

La qualité lexicale de textes produits à la fin du primaire et l'influence du contexte socioéconomique

Carole Boudreau
Université de Sherbrooke

Pascale Nootens
Université de Sherbrooke

Thierno M. O. Diallo
Australian Catholic University

Résumé

Cet article présente les résultats d'une étude visant à comparer des textes d'élèves de 6^e année du primaire provenant de milieux socioéconomiques contrastés. L'analyse des productions écrites a porté sur le choix lexical et les décisions concernant l'orthographe d'usage. Au total, 105 élèves (63 de milieux défavorisés et 42 de milieux favorisés) ont été soumis à une épreuve de production écrite d'une durée de cinq minutes, à partir d'une histoire séquentielle imagée. Les résultats révèlent des différences entre les milieux socioéconomiques à l'égard du pourcentage d'erreurs d'orthographe d'usage dans le texte, mais pas sur le plan de la diversité lexicale.

Mots-clés : production de textes, milieux défavorisés, variété lexicale, orthographe

Abstract

The article presents the results of a comparative study of texts produced by students from contrasting socio-economic backgrounds in the last stage of primary school. The analysis of students' written productions concentrates on lexical choices and spelling decisions in relation to standard spelling. A total of 105 students attending the sixth grade of primary school (63 from disadvantaged backgrounds, 42 from privileged backgrounds) completed a written production test based on a sequential illustrated narrative. The results reveal differences between socio-economic backgrounds in terms of the percentage of lexical spelling errors in written productions but not in terms of vocabulary diversity.

Keywords: text production, disadvantaged backgrounds, lexical variety, spelling

Introduction et problématique de la recherche

De nombreuses études ont mis en évidence un niveau de réussite scolaire plus faible chez les élèves de milieux défavorisés, en comparaison avec ceux des milieux favorisés, en particulier en lecture et en écriture (Barone, 2003-2004, Sévigny, 2003). Au Québec, une publication du ministère de l'Éducation (Ministère de l'Éducation du Québec, 2005) révèle qu'une proportion significativement plus élevée d'élèves de milieux défavorisés entre au secondaire avec un important retard¹. En 2008-2009, le taux de diplomation des élèves québécois était de 88 % en milieu très favorisé, comparativement à 69 % en milieu très défavorisé (Ministère de l'Éducation, du Loisir et du Sport, 2011). Ce rapport révèle également qu'une proportion moins élevée d'élèves de milieux défavorisés réussit l'épreuve d'écriture de 5^e secondaire, en comparaison à ceux des milieux favorisés. Plus le niveau socioéconomique est faible, plus le taux de réussite en écriture en 5^e secondaire est faible. Un vocabulaire restreint², un manque d'expérience lié à la littératie et peu d'occasions de vivre des expériences riches au plan culturel (Bradley & Corwyn, 2002; Brooks-Gunn, Rouse & McLanahan, 2007), peuvent expliquer la difficulté des élèves de milieux défavorisés à performer sur le plan scolaire, notamment en ce qui concerne les compétences à lire et à écrire. Le fait que les élèves des milieux défavorisés aient accès à un vocabulaire moins riche dans leur milieu familial (Dickinson, McCabe, Anastasopoulos, Peisner-Feinberg & Poe, 2003) crée une dissociation entre la langue utilisée en famille et celle valorisée par l'école (Romain, 2007). De plus, les élèves de milieux défavorisés ont des interactions moins riches avec l'écrit en milieu familial, alors qu'elles sont essentielles au développement de la conscience de l'écrit, elle-même essentielle à l'apprentissage formel de l'écrit (Snow, Burns & Griffin, 1998). Selon Hoff (2003), les enfants de familles défavorisées ont, par conséquent, des compétences linguistiques moins avancées que les enfants du même âge des milieux plus favorisés, notamment sur le plan du discours oral. Cette caractéristique linguistique des élèves de milieux défavorisés pourrait suggérer que certains éléments linguistiques oraux propres à ces élèves se retrouvent également dans leurs productions écrites.

1 Le retard est défini en tant qu'allongement du cheminement scolaire d'un à trois ans (Ministère de l'Éducation, du Loisir et du Sport, 2005).

2 Dans le cadre de cet article, le terme « lexicale » sera utilisé pour traiter de vocabulaire.

Jusqu'à maintenant, les études québécoises visant à comparer la réussite scolaire des élèves selon leur milieu socioéconomique ont principalement été effectuées auprès d'élèves du secondaire et se sont davantage consacrées au volet lecture qu'écriture. Aucun rapport ministériel, permettant de distinguer les élèves du primaire à l'égard de leur réussite en écriture en fonction de leur milieu socioéconomique, n'a été trouvé, d'où l'intérêt de se pencher sur la compétence à écrire d'élèves de fin du primaire issus de milieux défavorisés en comparaison à celle de leurs pairs de milieux plus favorisés. Cette étude entend donc poser un regard comparatif sur la qualité des textes produits par des élèves de fin du primaire de milieux socioéconomiques contrastés.

Cadre

La production de texte constitue une tâche complexe impliquant un coût cognitif, cela en raison des contraintes imposées par la capacité limitée de la mémoire et de l'attention, notamment chez l'apprenti scripteur (Fayol, 2007; Morin, Nootens, Labrecque & LeBlanc, 2009). Plusieurs études montrent d'ailleurs un lien entre le coût cognitif associé à certaines composantes rédactionnelles et à leur gestion, son fonctionnement en cours de rédaction et la qualité du texte produit (Bereiter & Scardamalia, 1987; Fayol, 2007; Olinghouse & Leaird, 2009). Dans cette perspective, plusieurs travaux ont montré que la fluidité du geste graphomoteur (Medwell, Strand & Wray, 2009), les compétences orthographiques ou plus globalement la fluidité en écriture (Graham, Berninger, Abbott, Abbott & Whitaker, 1997), les compétences syntaxiques (Beers & Nagy, 2009) et la qualité des choix lexicaux (Grobe, 1981; Deno, Marston & Mirkin, 1982; Olinghouse & Leaird, 2009) étaient corrélées à la qualité textuelle.

Cet article s'intéressera plus spécifiquement à deux composantes impliquées dans la qualité textuelle : l'orthographe d'usage et le lexique. L'orthographe d'usage correspond à la graphie correcte spécifique à chaque mot en lui-même, sans égard à sa fonction ou à son contexte d'apparition dans la phrase (Cogis, 2005; Lefrançois, Laurier, Lazure & Claing, 2008)³. Le lexique est pour sa part, tel que défini par Isaacson (1988), l'originalité et la maturité dans le choix des mots.

3 L'orthographe grammaticale ne sera pas abordée dans le cadre de cette contribution.

Orthographe d'usage, lexicale et qualité textuelle

Nombre de travaux se sont intéressés aux liens entre l'orthographe d'usage et la qualité textuelle. Soulignons d'emblée les imposantes recherches de Graham, Berninger et leurs collaborateurs (par ex., Graham et al., 1997; Berninger, Vaughan, Abbott, Brooks, Abbott, Rogan, et al., 1998; Berninger, Vaughan, Abbott, Begay, Coleman, Curtin, et al., 2002). Graham et al. (1997) ont exploré, plus globalement, les liens entre les capacités dites de bas niveau impliquées dans la transcription⁴ – soit le geste graphomoteur et l'orthographe d'usage – et la qualité textuelle de jeunes scripteurs. Les résultats de cette recherche ont montré que les composantes orthographique et graphomotrice combinées, liées à la fluidité de la transcription, contribuaient significativement dans la variance de la qualité des textes produits par les scripteurs du primaire. Dans le même sens, Graham (1999) a rapporté une corrélation positive importante – entre 40 et 50 % – entre les performances orthographiques et la qualité textuelle chez les scripteurs débutants. D'autres travaux (par ex., Berninger et al., 1998; Berninger et al., 2002), impliquant cette fois des programmes d'entraînement portant sur l'orthographe d'usage, ont quant à eux permis d'appuyer l'hypothèse de la relation causale entre performances orthographique et rédactionnelle. Allant en ce sens, les résultats de Berninger et al. (1998) ont montré une amélioration significative de la longueur des textes produits et des performances orthographiques en contexte rédactionnel chez des élèves de deuxième année ayant bénéficié d'un programme d'entraînement visant l'orthographe d'usage. Berninger et al. (2002) ont, pour leur part, rendu compte d'améliorations significatives des performances orthographiques et rédactionnelles chez des élèves en difficulté d'écriture de troisième année qui avaient bénéficié d'un double entraînement visant l'orthographe d'usage et la production écrite.

Les recherches concernant l'orthographe d'usage et son influence sur la qualité textuelle sont cependant essentiellement issues de l'anglais; elles se généralisent difficilement à l'orthographe française, en raison notamment de sa morphologie flexionnelle qui offre des défis particuliers à l'apprenti scripteur en contexte rédactionnel (Fayol, 2006). Pour cette raison, les travaux en français s'intéressant à la composante orthographique et à son implication lors de la transcription se sont surtout penchés sur les problématiques

4 La transcription porte sur la dimension langagière (lexique, syntaxe, agencements rhétoriques), mais implique également les dimensions orthographique et graphomotrice, qui ajoutent aux difficultés de transcription (Fayol, 2007).

liées à la gestion de l'orthographe dite grammaticale (Largy, 2005). Par ailleurs, ces travaux concernent, pour la plupart, les jeunes scripteurs de première à troisième année, plutôt que ceux de fin du primaire.

D'autres travaux ont exploré le lien entre le choix lexical et la qualité des textes produits, et se sont intéressés à la valeur prédictive du lexique, en comparaison avec d'autres variables – dont l'orthographe – potentiellement impliquées dans la qualité textuelle. Il n'existe toutefois pas de mesure unique du lexique : la justesse, la diversité et la maturité lexicale, de même que la longueur moyenne des mots ou de leurs syllabes, ont fait l'objet d'études corrélationnelles, qui tendent dans l'ensemble à montrer le lien entre ces différentes mesures et la qualité textuelle. Une étude de Grobe (1981), visant à explorer la valeur prédictive d'un ensemble de variables textuelles sur la qualité de textes narratifs rédigés par des élèves de cinquième année du primaire (N=950), de deuxième (N=1000) et de cinquième année du secondaire (N=750), a montré que les variables ayant trait au lexique expliquaient, dans le cadre de cette recherche, la plus grande variance dans la qualité textuelle. Les variables liées au lexique ont en effet permis d'expliquer une variance, respectivement de 81,4 %, 64,9 % et 59,8 %, dans la qualité des textes des élèves de cinquième année du primaire, de deuxième et de cinquième année du secondaire. Les résultats tendent par ailleurs à montrer que, parmi les variables liées au lexique, la diversité, soit le nombre de mots différents utilisés en production écrite, contribue pour la plus grande variance dans la qualité de la production d'élèves de secondaire cinq, et à la deuxième plus grande variance, après la variable orthographique, dans la qualité de la production écrite d'élèves de cinquième année.

D'autres recherches, comme celle de Deno, Marston & Mirkin (1982), ont permis de mettre en relief les liens entre la maturité lexicale – mesurée par le choix de mots moins fréquents⁵ – et la qualité des textes produits. L'étude de Deno et al. (1982), réalisée auprès de 130 élèves de 7 à 12 ans, dont 44 considérés en difficulté et 86 considérés normo-scripteurs, visait à mesurer la validité de prédicteurs de la qualité de l'expression écrite, telle la maturité lexicale, mais également la longueur moyenne des phrases ou propositions (*mean T-unit length*), le nombre total de mots, la longueur moyenne des mots

5 La maturité lexicale est construite par rapport au degré de sophistication du lexique utilisé dans un texte (Olinghouse & Leaird, 2009). Cette variable est mesurée en calculant le nombre ou la proportion de mots du texte ne faisant pas partie d'une liste de mots considérés fréquents ou usuels selon l'âge et le niveau scolaire d'une population d'élève (Deno et al., 1982).

ainsi que la variable orthographique (nombre de mots ou de séquences de lettres correctement écrits). Les résultats de cette étude montrent que la maturité lexicale, mesurée par la somme des mots du texte ne faisant pas partie d'une liste de référence de lexique usuel (Finn, 1972, cité dans Deno et al., 1982), est l'indice le plus fortement corrélé avec la qualité du texte, suivi de la composante orthographique.

Olinghouse & Leaird (2009) ont pour leur part cherché à identifier quelle variable de la composante lexicale – parmi la diversité et la maturité lexicales, la longueur moyenne des syllabes dans les mots produits et le nombre de mots polysyllabiques utilisés – pouvait expliquer la plus grande variance dans la qualité de textes narratifs. Pour ce faire, ils ont soumis, à trois moments dans l'année scolaire, des élèves de deuxième et de quatrième année (N=193) à une tâche de rédaction d'un court texte à partir d'une image. Les résultats rendent compte d'une corrélation significative entre les compétences orthographiques (ou la proportion de mots bien orthographiés) et la qualité des textes produits, de même qu'entre la diversité et la maturité lexicales – ou l'usage de mots peu fréquents – et la qualité textuelle, chez les élèves de deuxième et de quatrième année. Chez les premiers, la diversité lexicale s'est avérée le prédicteur le plus fiable de la qualité des textes, expliquant 20,6 % de la variance de la qualité textuelle. Chez les seconds, la diversité lexicale s'est avérée le second prédicteur le plus fiable de la qualité textuelle, derrière la longueur du texte, ces variables expliquant respectivement 5,7 % et 8,5 % de la variance de la qualité textuelle.

Une étude d'Olinghouse (Olinghouse & Wilson, 2013), réalisée cette fois auprès d'élèves de cinquième année (N=105), rend compte de résultats assez similaires quant au rôle du lexique dans la qualité de textes narratifs. Six mesures de la composante lexicale ont été réalisées dans le cadre de cette recherche: la diversité et la maturité lexicales, l'usage du lexique spécifique au domaine traité, le degré d'élaboration (par le calcul du nombre moyen de mots – déterminants, adjectifs et adverbes – permettant de modifier le nom ou le groupe nominal, et permettant de mesurer le degré d'élaboration du lexique), le registre de langue et l'usage du lexique dit académique. Les résultats montrent que seules la diversité et la maturité lexicales se sont avérées des prédicteurs fiables de la qualité des textes produits – expliquant 9 % de la variance de cette qualité – et que seule la diversité lexicale s'est avérée un fort prédicteur unique de la qualité textuelle, expliquant 8,4 % de cette variance. Ces résultats corroborent ceux obtenus par Olinghouse & Leaird (2009).

La présente recherche vise à mesurer les différences éventuelles entre les élèves de deux milieux socioéconomiques quant à la diversité lexicale et à la qualité de l'orthographe d'usage dans les productions écrites.

Méthodologie

Participants

Au total, 105 élèves de langue maternelle française en fin de 3^e cycle du primaire ont participé à l'étude, dont 59,4 % proviennent de milieux scolaires défavorisés (n= 60) alors que 40,6 % fréquentent une école en milieux favorisés (n=42). Les élèves ayant participé à l'étude sont répartis dans 34 classes appartenant à des écoles de la région de Québec, dont 13 en milieux défavorisés (indices de milieu socioéconomique 9 et 10⁶) et 21 en milieux favorisés (indices de milieu socioéconomique 3 et 4). Les milieux scolaires ont été approchés en fonction de leur indice de milieu socioéconomique et seuls les élèves pour lesquels les parents ont donné leur consentement ont participé à l'étude.

Instruments de mesure

Les habiletés de productions écrites des élèves ont été mesurées à partir du test *La production d'un texte écrit*, tiré de l'Épreuve des Compétences Linguistiques (ECL) (Khomsi, Nanty, Parbeau-Guéno & Pasquet, 2005). À partir d'une histoire séquentielle composée de six images, les élèves ont d'abord décrit chacune des images. Cinq minutes ont été allouées à cette tâche. En deuxième lieu, à partir de la même séquence d'images, les élèves ont rédigé une histoire, toujours dans un délai de cinq minutes. Avant de procéder à chacune des tâches, des consignes ont été transmises à l'ensemble des élèves de chaque classe ayant participé à l'étude. Les deux tâches de production ont été complétées individuellement. Seule la deuxième tâche, soit celle visant la rédaction d'une histoire, a

6 Les écoles du Québec sont classées sur une échelle allant de 1 à 10, le rang 1 étant considéré comme le plus favorisé; le rang 10 comme étant le plus défavorisé.

fait l'objet d'analyse scripturale. La *figure 1* présente les six images ayant été soumises aux deux tâches.

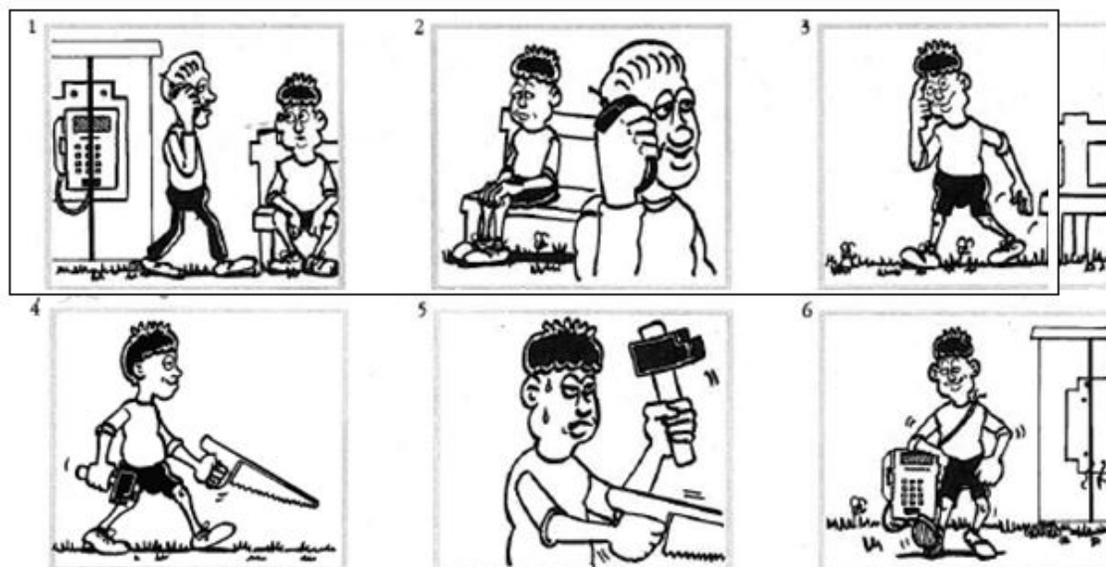


Figure 1. La production d'un texte écrit, tirée de l'Épreuve des Compétences Linguistiques au Collège (Khomsi et al., 2005)

Type de recherche et collecte de données

La présente recherche est de type exploratoire. La méthode dite *a posteriori* (Schriver, 1989), c'est-à-dire après la production du texte, a été utilisée pour effectuer l'analyse des textes. Le test visait à recueillir un échantillon de la production écrite des élèves et d'en faire une analyse sur le plan scriptural, notamment à l'égard de la diversité lexicale et de l'orthographe d'usage.

Analyse des données quantitatives

Un devis quantitatif a été utilisé pour mesurer la variété lexicale (Deno et al., 1982) et l'orthographe d'usage. La méthode du TTR (Type Token Ratio) (Retherford, 1993; Vermeer, 2000) a été retenue pour mesurer la diversité des mots exploités par le scripteur, celle-ci s'avérant la plus utilisée (Olinghouse & Leaird, 2009). Le ratio se calcule à partir du nombre de mots différents (uniques) sur le nombre total de mots dans le texte.

Un score élevé au TTR signifie une large diversité lexicale, comparativement à un score plus faible, qui suggère l'inverse. Le TTR permet de calculer à la fois la diversité lexicale pour différentes catégories de mots (noms, verbes, adverbes, déterminants, conjonctions, prépositions, adjectifs, pronoms), mais aussi pour l'ensemble du texte.

La variable liée à l'orthographe d'usage a été analysée à partir de la somme des erreurs d'orthographe d'usage dans les productions écrites, cela pour chacune des catégories de mots : noms, verbes, adverbes, déterminants, conjonctions, prépositions, adjectifs et pronoms. Cette somme a été calculée exclusivement à partir des mots erronés d'un point de vue lexical. À titre d'exemple, dans le segment de phrase « un enfants », une seule erreur est comptabilisée puisque la marque du pluriel en fin de mot relève de l'orthographe grammaticale.

Analyses statistiques

Pour répondre à la question de recherche, des analyses de corrélation ainsi que des analyses de variance (ANOVA) ont été effectuées à l'aide du logiciel SPSS 17. À la suite d'une première série d'analyses visant à distinguer les milieux socioéconomiques entre eux, huit variables ont été retenues pour la poursuite de l'analyse : les variables liées à la diversité des noms, des verbes, des adjectifs et des adverbes; le pourcentage de noms et d'adverbes; la somme des erreurs d'orthographe d'usage; et le TTR.

Résultats

Une première analyse descriptive (*Tableau 1*) selon le milieu socioéconomique permet de constater qu'à l'exception des variables liées au pourcentage de noms, d'adverbes et à la somme des erreurs d'orthographe d'usage, les résultats des élèves sont similaires d'un milieu socioéconomique à l'autre.

Tableau 1. Statistiques descriptives des variables de l'étude

Variables	Milieu favorisé (N=42)			Milieu défavorisé (N=60)		
	Min/Max	Moy.	É-T	Min/Max	Moy.	É-T
Diversité des noms	2/12	6,95	2,13	3/12	6,55	1,91
Diversité des verbes	5/13	8,31	1,72	4/13	8,22	1,88
Diversité des adjectifs	0/6	1,90	1,65	0/6	1,87	1,35
Diversité des adverbes	0/6	2,14	1,65	0/5	1,58	1,21
Pourcentage noms	13,56/26,92	19,65	3,29	15,25/30	21,06	3,26
Pourcentage adverbes	0/18,18	5,2	3,81	0/13,33	3,85	2,89
Somme des erreurs*	0/7	2,52	1,98	0/14	3,85	2,89
TTR	2,86/11,9	8,40	2,08	4,43/11,9	7,78	1,74

* Il s'agit des erreurs d'orthographe d'usage.

Analyse des corrélations

La matrice de corrélations du *Tableau 2* illustre les relations entre les variables de l'étude. La diversité des noms est significativement corrélée avec la diversité des verbes, $r=0,4$, la diversité des adjectifs, $r=0,46$, la diversité des adverbes, $r=0,34$, et le TTR, $r=0,83$, pour le milieu favorisé. Pour le milieu défavorisé, la diversité des noms est significativement corrélée avec l'ensemble des variables de l'étude (avec des coefficients de corrélation positifs variant entre 0,31 et 0,81), sauf avec le pourcentage de noms dans le texte et la somme des erreurs d'orthographe d'usage.

La diversité des verbes est significativement corrélée avec le pourcentage de noms dans le texte, $r=0,35$, et le TTR, $r=0,67$, pour le milieu favorisé. Pour le milieu défavorisé, la diversité des verbes est significativement corrélée avec l'ensemble des variables de l'étude (avec des coefficients de corrélation $r=0,39$ pour la diversité des noms, de $r=0,31$ pour la diversité des adverbes, de $r=0,3$ pour le pourcentage de noms dans le texte, de $r=0,3$ pour le pourcentage des adverbes dans le texte, de $r=0,26$ pour la somme des erreurs, et de $r=0,71$ pour le TTR) sauf avec la diversité des adjectifs.

La diversité des adjectifs est significativement corrélée avec le TTR, $r=0,43$, pour le milieu favorisé. Pour le milieu défavorisé, la diversité des adjectifs est corrélée avec la diversité des adverbes, $r=0,38$, le pourcentage des adverbes dans le texte, $r=0,28$, et le TTR, $r=0,51$.

La diversité des adverbes est significativement corrélée avec le pourcentage des noms dans le texte, $r=0,33$, le pourcentage des adverbes, $r=0,87$, et le TTR, $r=0,57$, pour le milieu favorisé. Pour le milieu défavorisé, la diversité des adverbes est significativement corrélée avec le pourcentage des adverbes dans le texte, $r=0,95$, et le TTR, $r=0,62$.

Le pourcentage des noms dans le texte est significativement corrélé avec la somme des erreurs, $r=0,34$, pour le milieu défavorisé. Pour les deux milieux, le pourcentage des adverbes dans le texte est significativement corrélé avec le TTR (avec un coefficient de corrélation r de $0,32$ pour le milieu favorisé et de $0,5$ pour le milieu défavorisé). La somme des erreurs ayant trait à l'orthographe d'usage est significativement corrélée avec le TTR, $r=0,29$, pour le milieu défavorisé.

En bref, la diversité lexicale est corrélée significativement avec la diversité des noms, des verbes, des adjectifs et des adverbes, cela dans les deux milieux socioéconomiques. Par ailleurs, il n'y a pas de corrélations significatives entre la somme des erreurs d'orthographe lexicale et les autres variables à l'étude en milieu favorisé. En milieu défavorisé, toutefois, la somme des erreurs est corrélée à la diversité des adverbes et au pourcentage de noms utilisés.

Tableau 2. Matrice de corrélations des variables de l'étude par milieu socioéconomique

Variables	1	2	3	4	5	6	7	8
1. Diversité des noms	-	0,40**	0,46**	0,34*	0,26	0,22	0,16	0,83**
2. Diversité des verbes	0,39**	-	0,05	0,25	-0,35*	0,02	0,13	0,67**
3. Diversité des adjectifs	0,4**	0,1	-	0,08	0,17	-0,12	-0,09	0,43**
4. Diversité des adverbes	0,45**	0,31*	0,38**	-	-0,33*	0,87**	0,22	0,57**
5. Pourcentage noms	0,14	-0,3*	-0,05	-0,14	-	-0,14	-0,22	-0,16
6. Pourcentage adverbes	0,31*	0,3*	0,28*	0,95**	-0,16	-	0,22	0,32*
7. Somme des erreurs	0,17	0,26*	0	0,19	-0,34*	0,21	-	0,27
8. TTR	0,81**	0,71**	0,51**	0,62**	-0,2	0,5**	0,29*	-

Note. Les coefficients sous la diagonale sont pour le milieu défavorisé; ceux au-dessus de la diagonale sont pour le milieu favorisé.

* $p<0,05$; ** $p<0,01$.

Analyse de variance

Les résultats de l'analyse de variance sont présentés au *Tableau 3*. Ces résultats indiquent que les élèves des deux milieux se différencient sur le pourcentage des noms utilisés dans le texte ($F(1,100)=48,81$ $p<0,05$), le pourcentage des adverbes utilisés dans le texte ($F(1,100)=45,35$ $p<0,05$) et sur la somme des erreurs ($F(1,100)=43,45$ $p<0,05$). Les élèves des milieux défavorisés ont un pourcentage de noms utilisés dans le texte plus élevé ($M=21,06$) que ceux de milieux favorisés ($M=19,65$) et les premiers présentent plus d'erreurs d'orthographe d'usage dans leur copie ($M=3,85$) que les seconds ($M=2,52$). Par ailleurs, les élèves de milieux favorisés ont un pourcentage des adverbes utilisés dans le texte ($M=5,2$) plus élevé que ceux des milieux défavorisés ($M=3,85$). Les deux milieux ne se différencient pas sur la diversité des noms, des verbes, des adjectifs, des adverbes ainsi que sur le TTR.

Tableau 3. Résultats de l'analyse de variance des variables de l'étude

Variables	Somme des carrés de type III	dl	F	p
Diversité des noms	4	1	0,99	0,32
Diversité des verbes	0,21	1	0,07	0,80
Diversité des adjectifs	0,04	1	0,02	0,89
Diversité des adverbes	7,74	1	3,91	0,05
Pourcentage noms	48,81	1	4,57	0,04
Pourcentage adverbes	45,35	1	4,17	0,04
Somme des erreurs	43,45	1	6,66	0,01
TTR	9,47	1	2,66	0,11

Discussion

Cette étude vise à mesurer ce qui distingue les élèves de fin du primaire de milieux défavorisés de ceux des milieux favorisés sur le plan de la réussite en écriture. Considérant la complexité et la diversité des critères d'évaluation de la qualité d'un texte écrit en contexte scolaire, de même que l'état des recherches montrant que les compétences orthographiques et la diversité lexicale expliquent une part importante de la variance dans la qualité textuelle, l'analyse a porté sur ces composantes.

La diversité lexicale

Les résultats de l'étude montrent que les élèves de milieux socioéconomiques défavorisés et favorisés ne se distinguent pas entre eux, du moins globalement, sur le plan de la diversité lexicale, que celle-ci soit mesurée par catégorie (noms-verbs-adverbes-adjectifs) ou pour l'ensemble du texte. Ainsi, les élèves des deux milieux socioéconomiques utilisent une diversité lexicale comparable. Bien qu'une analyse qualitative de cette variable préalablement effectuée suggère un lexique plus riche en milieux favorisés (par exemple, le mot cellulaire est utilisé plus fréquemment en milieux favorisés qu'en milieux défavorisés), les résultats de l'analyse quantitative ne vont pas dans ce sens. Ces résultats ne corroborent pas, à première vue, ceux qui ressortent de travaux antérieurs s'étant intéressés à l'effet de la pauvreté sur le développement lexical chez l'enfant. Tabors, Snow & Dickinson (2001) soulignent en effet l'impact néfaste majeur de la pauvreté sur l'étendue et la qualité du lexique de l'enfant lors de son entrée dans l'écrit. Toutefois, les élèves sélectionnés pour notre recherche sont à la fin de leur parcours primaire, non pas au stade de l'entrée dans l'écrit. Il est donc possible de supposer un effet positif de la scolarisation sur le développement lexical des élèves de milieux socioéconomiques défavorisés.

Si, tel que vu, les élèves à l'échantillon ne se distinguent pas selon leur milieu socioéconomique sur le plan de la diversité lexicale de façon globale, les résultats de cette étude révèlent cependant que les élèves de milieux défavorisés tendent à utiliser, de manière significative, un nombre plus élevé de noms dans leur production écrite, que ces noms soient ou non récurrents, que le font ceux des milieux favorisés. Les analyses de corrélation montrent par ailleurs qu'en milieux défavorisés, le pourcentage de noms utilisés dans le texte ne corrèle pas avec la diversité des noms. Ainsi, si les élèves de ces milieux utilisent davantage de noms dans leur texte, ceux-ci sont peu variés. De leur côté, les élèves de milieux favorisés tendent à utiliser davantage d'adverbes dans leur texte que ceux des milieux défavorisés. Les analyses révèlent par ailleurs une corrélation positive entre le pourcentage d'adverbes dans le texte et la diversité des adverbes, et ce, peu importe le milieu socioéconomique. Ainsi, plus les élèves utilisent d'adverbes dans leur texte, plus ceux-ci sont variés. Notons cependant que la diversité lexicale liée aux adverbes n'est pas significativement différente entre les milieux, seul le pourcentage d'adverbes dans le texte permet de distinguer de façon significative les deux milieux.

Considérant les fonctions différentes qu'occupent les adverbes et les noms dans une production verbale écrite, il est à propos de se questionner sur le fait que les élèves de milieux défavorisés utilisent davantage de noms que ceux de milieux favorisés, alors que les élèves de milieux favorisés utilisent, quant à eux, plus d'adverbes. Hudson (2009) s'est d'ailleurs intéressé à l'utilisation relative des diverses catégories de mots en production écrite chez des apprentis scripteurs et des scripteurs experts; l'auteur s'est également questionné quant aux liens éventuels entre la proportion de mots de chaque catégorie exploitée par le scripteur et la maturité textuelle. L'ensemble des données recueillies sur la question tend à montrer, selon Hudson (2009), que la maturité textuelle est avant tout intimement associée à la diversité lexicale. Ainsi, la proportion de noms ou d'adverbes d'un texte semble devoir être considérée avec prudence en tant que prédicteur – notamment unique – de la qualité textuelle.

D'autres mesures de la qualité, ou maturité lexicale, prenant en compte les catégories de mots ont d'ailleurs été explorées. Wells (1960) a cherché à mesurer la maturité lexicale de textes de scripteurs adultes en tenant compte, non pas des proportions relatives des diverses catégories de mots au sein des textes produits, mais bien de la proportion de certains groupements de catégories lexicales. L'auteur a ainsi mis en relief l'utilisation préférentielle, respectivement, d'un « style nominal », caractérisé par l'exploitation de groupes de noms communs et propres accompagnés d'adjectifs et de prépositions; ou plutôt d'un « style verbal », exploitant davantage de groupes formés de verbes, d'adverbes et de pronoms dans des textes de différents genres rédigés par des scripteurs adultes. L'auteur a ainsi relevé la prévalence du « style verbal » – associant des groupes de verbes, adverbes et pronoms – dans les textes qualifiés « d'imaginatifs » (auxquels on peut associer le texte narratif) produits par les scripteurs qualifiés d'experts. Peut-on suggérer que, si ces scripteurs plus experts usent préférentiellement du style verbal en contexte de rédaction de texte narratif, l'exploitation de ce type de groupements lexicaux pourrait être liée, de quelque façon, à une plus grande maturité lexicale chez les apprentis scripteurs de notre échantillon? Ce lien reste à explorer, bien que – instinctivement – l'exploitation plus importante de catégories de mots – tels les adverbes et pronoms, notamment autres que personnels – paraisse rendre compte d'un développement plus important des connaissances lexicales.

L'orthographe d'usage

Les résultats de l'étude montrent une différence significative entre les élèves de milieux défavorisés en faveur de ceux des milieux favorisés, sur le plan de la somme des erreurs d'orthographe d'usage contenues dans le texte. L'analyse comparative de la proportion des erreurs commises en fonction du milieu socioéconomique ne permet toutefois pas de mieux comprendre les éventuels distinctions ou lieux communs entre les deux groupes au regard des performances en orthographe d'usage. Un répertoire des différentes graphies utilisées dans chaque milieu et une analyse globale qualitative des erreurs semblent ainsi nécessaires pour caractériser plus finement les différences entre les milieux socioéconomiques, à cet égard, au sein de l'échantillon. Ceci fera l'objet d'une recherche ultérieure.

Par ailleurs, la recherche s'est limitée à la mesure des performances en orthographe d'usage chez les élèves de milieux socioéconomiques contrastés. Or, pour l'apprenti scripteur qui se doit de maîtriser le français écrit, la morphologie et les règles de genre et de nombre posent de grands défis, notamment à cause de ses marques grammaticales souvent inaudibles (Pacton & Fayol, 2004; Laparra, 2010). Des recherches menées auprès d'élèves francophones de milieux socioéconomiques contrastés, et tenant compte des liens entre les composantes orthographiques – orthographe d'usage et grammaticale – et la qualité textuelle, devraient permettre de poser un regard comparatif plus complet à cet égard entre les milieux.

Il semble exister, à fin de comparaison de nos résultats, peu de données antérieures traitant d'éventuelles distinctions et lieux communs entre milieux socioéconomiques sur le plan des compétences quant à l'orthographe d'usage en contexte rédactionnel. Toutefois, l'implication du milieu socioéconomique dans le développement de certaines compétences et connaissances de base liées au développement ultérieur des compétences orthographiques a fait l'objet de certaines études. Dans ce sens, Billard et ses collègues (2010) se sont intéressés aux facteurs prédictifs, dont le milieu socioéconomique⁷, des compétences en orthographe, mais aussi en lecture, auprès d'élèves (N=181) évoluant entre la deuxième et la quatrième année du primaire. Les élèves ont été scindés

7 Plus spécifiquement, les auteurs ont ciblé de faibles lecteurs de quartiers défavorisés, bilingues ou pas, et ont cherché à étudier les facteurs prédictifs des compétences en lecture et en orthographe en tenant compte de l'influence mutuelle de facteurs cognitifs, comportementaux et socioéconomiques.

en trois groupes sur la base de leur niveau initial de compétence en lecture. Les résultats montrent qu'aucune variable seule liée au milieu socioéconomique (notamment la profession des parents et leur niveau de scolarité) ne s'est avérée – lors de cette étude – liée à un groupe avec évolution très défavorable en termes de compétences à l'écrit. Une faible compétence en conscience phonologique a, en revanche, été le facteur le plus clairement associé à un très faible niveau en lecture, peu importe le milieu socioéconomique d'origine. La conscience phonologique est également le facteur le plus prédictif des compétences en orthographe, comme le rappellent Billard et al. (2010). Or, plusieurs travaux, dont ceux de Duncan & Seymour (2000) et de Molfese, Modglin, Beswick, Neeamon et al. (2006), tendent à montrer qu'au préscolaire, les enfants de milieux socioéconomiques défavorisés sont plus faibles que leurs pairs de milieux plus favorisés en conscience phonologique, dont l'émergence est elle-même, comme mentionné, favorisée par le développement du lexique (Gotswami, 2001); dans les milieux défavorisés, le lexique auquel l'enfant a accès, notamment dans la famille – premier lieu d'acquisition des connaissances lexicales – est en revanche souvent moins riche (Dickinson et al., 2003). Aussi, au terme de leur recherche, Billard et al. (2010) considèrent que les liens entre les facteurs socioéconomiques et l'évolution des compétences liées à la lecture et à l'orthographe chez l'élève sont complexes. Selon eux, les conditions de vie, tels la pauvreté et un environnement source de difficulté pour l'enfant, ont des effets néfastes sur le développement des compétences à l'écrit, mais que cet effet est plus délétère que causal.

Du reste, nombre de travaux – dont ceux de Graham (1999), de Berninger et al. (2002) – rendent compte du lien entre les compétences ayant trait à l'orthographe d'usage et la qualité textuelle. Si différents processus et connaissances linguistiques peuvent être mis à contribution pour permettre au scripteur d'orthographier un mot – conscience phonémique, connaissances liées au principe alphabétique et à certains patrons orthographiques, connaissances sémantiques, conscience morphologique (Masterson & Apel, 2010) – les connaissances stockées en mémoire, ayant trait à la représentation orthographique spécifique des mots, permettraient au scripteur un accès direct à ces mots en contexte de rédaction, ceci libérant d'autant la mémoire de travail, dont la capacité limitée pourrait dès lors être dédiée aux processus de plus haut niveau (Berninger et al., 2002; Masterson & Apel, 2010; Ehri, 2000). En somme, si tel que l'ont proposé certains auteurs (Duncan & Seymour, 2000; Molfese et al., 2006), le développement de la conscience phonologique s'avère souvent déficitaire chez les enfants d'âge préscolaire

de milieux socioéconomiques défavorisés, et que cette variable constitue par ailleurs le plus fort prédicteur des compétences ultérieures en orthographe d'usage (Billard et al., 2010), il serait logique de poser l'hypothèse que chez les apprentis scripteurs de milieux plus défavorisés, l'orthographe pose plus de défis et occupe ainsi une plus grande part de l'attention du scripteur. En tenant compte des limites de capacité de la mémoire de travail (Ehri, 2000), ces apprentis scripteurs montreraient, d'une part, des performances moindres sur le plan de l'orthographe, ce qui s'est avéré le cas dans la présente étude en ce qui a trait à l'orthographe d'usage. D'autre part, étant donné cette plus grande charge cognitive dédiée aux choix orthographiques entourant les mots produits, il paraît probable qu'une charge cognitive moindre puisse être dédiée, chez ces scripteurs, aux processus de haut niveau tels que la génération des idées, l'organisation textuelle, la syntaxe et le choix lexical, tout ceci influant sur la qualité du texte produit. Ainsi, dans le cadre de cette recherche, certaines distinctions observées entre milieux socioéconomiques contrastés au regard du choix lexical – notamment l'utilisation plus importante et plus diversifiée des adverbes en milieux favorisés, mots qui seraient liés à une plus grande maturité lexicale (Wells, 1960) – pourraient-elles être à la fois dues à l'état des connaissances orthographiques et lexicales des apprentis scripteurs de milieux favorisés, mais aussi, en partie, au fait que ces mêmes scripteurs aient pu dédier, en général, plus de charges cognitives au choix de mots plus « recherchés », si certaines compétences de base – dont les compétences orthographiques – sont plus automatisées? Cette hypothèse paraît plausible, mais devra être explorée pour ce qui est d'apprentis scripteurs francophones de milieux socioéconomiques contrastés.

Conclusion

Cette étude, dont l'objet visait à mieux comprendre ce qui différencie la réussite en écriture des élèves de milieux défavorisés de ceux de milieux favorisés, a porté essentiellement sur deux composantes liées à la qualité d'une production écrite, soit la diversité lexicale et l'orthographe d'usage. Plusieurs études montrent une différence entre les élèves de milieux socioéconomiques différents quant à la réussite en écriture. Les résultats obtenus lors de cette étude vont dans le même sens, puisqu'ils révèlent que les élèves de milieux socioéconomiques favorisés se distinguent au détriment de leurs pairs des

milieux défavorisés, sur le plan de la somme des erreurs d'orthographe d'usage produites dans le texte, tout comme sur le plan de l'utilisation plus importante d'adverbes, souvent associée à une plus grande maturité lexicale (Wells, 1960). Ce sont toutefois sur ces seuls éléments qu'il a été possible de distinguer les élèves selon leurs milieux respectifs. Néanmoins, vu l'influence – reconnue par plusieurs chercheurs – de la maturité des choix lexicaux et du nombre d'erreurs orthographiques sur le jugement de la qualité textuelle, l'hypothèse qu'un biais négatif à l'égard des productions textuelles des élèves de milieux défavorisés – et par conséquent de leurs compétences rédactionnelles – puisse se poser, pourrait constituer une partie de l'écart entre milieux contrastés au regard de la réussite en écriture. Les résultats de la présente étude permettent néanmoins de mieux comprendre certains aspects de la problématique liés à la réussite en écriture des élèves de milieux défavorisés. Par ailleurs, l'ajout de commentaires métagraphiques à la suite des productions écrites, lors d'une prochaine étude, permettrait de documenter davantage la démarche cognitive des élèves au moment de la tâche (Fayol, 2007), et par le fait même, de mieux comprendre certains choix lexicaux, notamment les choix adverbiaux et nominaux.

Références

- Barone, D. (2003–2004). Second grade is important: Literacy instruction and learning of young children in a high-poverty school. *Journal of Literacy Research*, 35, 965–1019.
- Beers, S. F. & Nagy, W. E. (2009). Syntactic complexity as a predictor of adolescent writing quality: Which measures? Which genre? *Reading and Writing*, 22, 185–200.
- Bereiter, C. & Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Berninger, V. W., Vaughan, K. B., Abbott, R. D., Brooks, A., Abbott, S. P., Rogan, A., Reed, E., & Graham, S. (1998). Early intervention for spelling problems: Teaching spelling units of varying size within a multiple connection framework. *Journal of Educational Psychology*, 90, 587–605.
- Berninger, V. W., Vaughan, K., Abbott, R. D., Begay, K., Coleman, K. B., Curtin, G., Hawkins, J. M., & Graham, S. (2002). Teaching spelling and composition alone and together: Implications for the simple view of writing. *Journal of Educational Psychology*, 94(2), 291–304.
- Billard, C., Bricout, L., Ducot, G., Richard, J., Ziegler, J., & Fluss, J. (2010). Évolution des compétences en lecture, compréhension et orthographe en environnement socioéconomique défavorisé et impact des facteurs cognitifs et comportementaux sur le devenir à deux ans. *Revue d'Épidémiologie et de Santé Publique*, 58, 101–110.
- Bradley, R. H. & Corwyn, R. F. (2002). Socioeconomic status and child development. *Annual Review Psychologic*, 53, 371–399.
- Brooks-Gunn, J., Rouse, C. E., & McLanahan, S. (2007). Racial and ethnic gaps in school readiness. Dans R. C. Pianta, M. J. Cox & K. L. Snow (dir.), *School readiness and the transition to kindergarten in the era of accountability* (p. 283–306). Baltimore, MD: Brooks.
- Cogis, D. (2005). *Pour enseigner et apprendre l'orthographe*. Paris, FR: Delagrave.

- Deno, S., Marston, D., & Mirkin, P. (1982). Valid measurement procedures for continuous evaluation of written expression. *Exceptional Children*, 48, 368–371.
- Dickinson, D., McCabe, A., Anastasopoulos, L., Peisner-Feinberg, E., & Poe, M. (2003). The comprehensive language approach to early literacy: The interrelationships among vocabulary, phonological sensitivity, and print knowledge among preschool-aged children. *Journal of Educational Psychology*, 95(3), 465–481.
- Duncan, L. G. & Seymour, P. H. K. (2000). Socioeconomic differences in foundation-level literacy. *British Journal of Psychology*, 91, 145–166.
- Ehri, L. (2000). Learning to read and learning to spell: Two sides of a coin. *Topics in Language Disorders*, 20(3), 19–49.
- Fayol, M. (2006). L'orthographe et son apprentissage. *Les journées de l'ONL. Enseigner la langue : orthographe et grammaire* (p. 53–74). Paris, FR: Observatoire National de la Lecture.
- Fayol, M. (2007). La production de textes et son apprentissage. *Les journées de l'ONL. Écrire des textes, l'apprentissage et le plaisir* (p. 21–34). Paris, FR: Observatoire National de la Lecture.
- Gotswami, U. (2001). Early phonological development and the acquisition of literacy. Dans S.B. Neuman & D.K. Dicksison (dir.), *Handbook of early literacy research* (p. 111–125). New York, NY: Guilford Press.
- Graham, S. (1999). Handwriting and spelling instruction for students with learning disabilities: A review. *Learning Disabilities Quarterly*, 22, 78–79.
- Graham, S., Berninger, V. W., Abbott, R. D., Abbott, S. P., & Whitaker, D. (1997). Role of mechanics in composing of elementary school students: A new methodological approach. *Journal of Educational Psychology*, 89(1), 170–182.
- Grobe, G. (1981). Syntactic maturity, mechanics, and vocabulary as predictors of quality ratings. *Research in the Teaching of English*, 15(1), 75–85.
- Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development*, 74, 1368–1378.

- Hudson, R. (2009). Measuring maturity. Dans R. Beard, D. Myhill, J. Riley, & M. Nystrand (dir.), *The Sage handbook of writing development* (p. 349–361). Thousand Oaks, CA: Sage Publications.
- Isaacson, S. (1988). Assessing the writing product: Qualitative and quantitative measures. *Exceptional Children*, 54, 528–534.
- Khomsi, A., Nanty, I., Parbeau-Guéno, A., & Pasquet, F. (2005). *Évaluation des compétences linguistiques au collège*. Paris, FR: Les éditions du centre de psychologie appliquée.
- Laparra, M. (2010). Pour un enseignement progressif de l'orthographe dite grammaticale du français. *Repères*, 41, 35–46.
- Largy, P. (2005). Approche psycholinguistique de la mise en œuvre de la morphologie flexionnelle du nombre verbal en français langue maternelle. Dans J. Granfeldt & S. Schlyter (dir.), *Acquisition et production de la morphologie flexionnelle* (p. 89–96), PERLES, 20, Université de Lund.
- Lefrançois, P., Laurier, D., Lazure, M., & Claing, R. (2008). *Évaluation de l'efficacité des mesures visant l'amélioration du français écrit du primaire à l'université*. Montréal, QC: Office québécois de la langue française.
- Masterson, J. & Apel, K. (2010). Linking characteristics discovered in spelling assessment to intervention goals and methods. *Learning Disabilities Quarterly*, 33(3), 185–198.
- Medwell, J., Strand, S., & Wray, D. (2009). The links between handwriting and composing for Y6 children. *Cambridge Journal of Education*, 39(3), 329–344.
- Ministère de l'Éducation du Québec (2005). *La réussite scolaire des garçons et des filles. L'influence du milieu socioéconomique*. Québec, QC: Gouvernement du Québec.
- Ministère de l'Éducation, du Loisir et du Sport (2011). *Agir autrement. La stratégie d'intervention Agir Autrement (SIAA). Contrer les écarts de réussite entre les milieux défavorisés et ceux qui sont plus favorisés*. Québec, QC: Gouvernement du Québec.
- Molfese, V. J., Modglin, A. A., Beswick, J. L., Neamon, J. D., Berg, S. A., Berg, C. J., & Molnar, A. (2006). Letter knowledge, phonological processing, and print

- knowledge: Skill development in nonreading preschool children. *Journal of Learning Disabilities*, 39, 296–305.
- Morin, M.-F., Nootens, P., Labrecque, A.-M., & LeBlanc, I. (2009). *Synthèse de connaissances sur l'enseignement de l'écriture à l'école primaire*. Production écrite réalisée dans le cadre d'un contrat de recherche pour le MELS.
- Olinghouse, N. G. & Leaird, J. T. (2009). The relationship between measures of vocabulary and narrative writing quality in second - and fourth - grade students. *Reading and Writing*, 22, 545–565.
- Olinghouse, N. J. & Wilson, J. (2013). The relationship between vocabulary and writing quality in three genres. *Reading and Writing*, 26, 45–65.
- Pacton, S. & Fayol, M. (2004) Learning to spell a deep orthography. The case of French. Dans Bernan, R. & Gillis, T. (dir.) *Trends in language acquisition research* (p. 164–176). Dordrecht, NL: Kluwer.
- Retherford, K. (1993). *Guide to analysis of language transcripts* (2e éd.). Eau Claire, WI: Thinking Publications.
- Romain, C. (2007). L'emploi des temps et des organisateurs textuels dans des textes narratifs d'élèves de 9 à 14 ans issus de milieux socioculturels contrastés. *Revue des sciences de l'éducation*, 33(1), 209–235.
- Schraver, K. A. (1989). Evaluating text quality: The continuum from text-focused to reader-focused methods. *Transactions on Professional communication*, 32(4), 238–255.
- Sévigny, D. (2003). *Impact de la défavorisation socioéconomique sur la diplomation des élèves inscrits dans les écoles secondaires publiques de l'Île de Montréal*. Montréal, QC: Comité de gestion de la taxe scolaire de l'Île de Montréal.
- Snow, C. E., Burns, M. S., & Griffin, P. (1998). *Preventing Reading Difficulties in Young Children*. Washington, DC: National Academy Press.
- Tabors, P., Snow, C. E., & Dickinson, D. (2001). Homes and schools together : Supporting language and literacy development. Dans Dickinson, D. K. & Tabors, P. O. (dir.), *Beginning literacy with language* (p. 313–334). Baltimore, MD: Brookes Publishing.

Vermeer, A. (2000). Coming to grips with lexical richness in spontaneous speech data.

Language Testing, 17(1), 65–84.

Wells, R. (1960). Nominal and verbal style. Dans Sebeok, T. A. (dir.), *Style in language*

(p. 213–220). Cambridge, MA: The MIT Press.