

Direction des bibliothèques

AVIS

Ce document a été numérisé par la Division de la gestion des documents et des archives de l'Université de Montréal.

L'auteur a autorisé l'Université de Montréal à reproduire et diffuser, en totalité ou en partie, par quelque moyen que ce soit et sur quelque support que ce soit, et exclusivement à des fins non lucratives d'enseignement et de recherche, des copies de ce mémoire ou de cette thèse.

L'auteur et les coauteurs le cas échéant conservent la propriété du droit d'auteur et des droits moraux qui protègent ce document. Ni la thèse ou le mémoire, ni des extraits substantiels de ce document, ne doivent être imprimés ou autrement reproduits sans l'autorisation de l'auteur.

Afin de se conformer à la Loi canadienne sur la protection des renseignements personnels, quelques formulaires secondaires, coordonnées ou signatures intégrées au texte ont pu être enlevés de ce document. Bien que cela ait pu affecter la pagination, il n'y a aucun contenu manquant.

NOTICE

This document was digitized by the Records Management & Archives Division of Université de Montréal.

The author of this thesis or dissertation has granted a nonexclusive license allowing Université de Montréal to reproduce and publish the document, in part or in whole, and in any format, solely for noncommercial educational and research purposes.

The author and co-authors if applicable retain copyright ownership and moral rights in this document. Neither the whole thesis or dissertation, nor substantial extracts from it, may be printed or otherwise reproduced without the author's permission.

In compliance with the Canadian Privacy Act some supporting forms, contact information or signatures may have been removed from the document. While this may affect the document page count, it does not represent any loss of content from the document.

Université de Montréal

L'apport des correcticiels
pour la correction de textes d'élèves du secondaire

par
Marie-Hélène Mireault

M.A. Mesure et évaluation en éducation
Sciences de l'éducation

Mémoire présenté à la Faculté des études supérieures
en vue de l'obtention de grade M.A
en Mesure et évaluation en éducation

Avril 2009

© Marie-Hélène Mireault, 2009



Université de Montréal
Faculté des études supérieures

Ce mémoire intitulé :

L'apport des correcticiels
pour la correction de textes d'élèves du secondaire

présenté par :

Marie-Hélène Mireault

a été évalué par un jury composé des personnes suivantes :

Michel Laurier
président du jury

Jean-Guy Blais
directeur de recherche

Marie-Claude Boivin
codirectrice

Pascale Lefrançois
membre du jury

Résumé

L'objectif de cette recherche était de comparer la correction de la qualité de la langue effectuée par des correcteurs humains à celle effectuée par un correcticiel lors d'une épreuve de français au secondaire. Nous avons voulu mettre en parallèle les résultats des deux types de correction afin d'établir l'existence d'éventuelles différences entre la correction humaine et la correction automatisée. De plus, nous voulions juger de l'utilité du correcticiel *Antidote* lors des épreuves de français. Pour ce faire, nous avons appliqué le correcticiel *Antidote* sur un corpus constitué de 45 textes d'élèves de deuxième année du premier cycle du secondaire qui avaient déjà été corrigés par des correcteurs du MELS (ministère de l'Éducation, du Loisir et du Sport). L'analyse quantitative des données nous a permis de constater que plus le niveau du texte est élevé, plus la correction d'*Antidote*, en terme de nombre d'erreurs détectées, se rapproche de la correction effectuée par les correcteurs humains. L'analyse qualitative, quant à elle, nous a montré les points forts d'*Antidote*, soit la détection des erreurs d'orthographe d'usage plus constante que chez les correcteurs humains et l'analyse plus détaillée des erreurs, ainsi que ses points faibles, qui touchent principalement le domaine de la syntaxe. De nombreuses différences ont également été observées quant à la classification de certaines erreurs.

Mots clés : correcticiel, correction, évaluation, ordinateur, production écrite, *Antidote*

Abstract

The objective of this research was to compare the marking of the language accuracy made by human raters with the one made by a grammar-spell checker on a secondary level French essay. We wanted to compare the results from the two types of marking to establish the existence of potential differences between human marking and automated marking. We also wanted to determine the usefulness of the grammar-spell checker *Antidote* in marking French exams. To do this, we applied the grammar-spell checker *Antidote* to a sample of 45 Grade 8 student essays already marked by MELS' (ministère de l'Éducation, du Loisir et du Sport) raters. The quantitative analysis of our data showed us that the higher the quality of the text, the more *Antidote*'s marking, in terms of errors found, was closer to that of the marking made by its human counterparts. The qualitative analysis, for its part, showed *Antidote*'s strengths and weaknesses: a greater constancy than human scorers in the detection of spelling mistakes and a more detailed analysis of errors in general, but an important deficiency concerning syntax. Numerous differences were equally observed with regards to the classification of certain errors.

Key words: grammar-spell checker, correction, evaluation, computer, essays, *Antidote*.

Table des matières

Identification du jury.....	ii
Résumé.....	iii
Abstract.....	iv
Table des matières.....	v
Liste des tableaux.....	viii
Liste des figures.....	ix
Liste des sigles et abréviations.....	x
Remerciements.....	xii
Introduction.....	1
Premier chapitre – Problématique de la recherche.....	3
1.1 Utilisation des ordinateurs dans les écoles.....	3
1.2 Utilisation des correcticiels aux États-Unis	4
1.3 Avantages des correcticiels.....	5
1.4 Avantages possibles de l'utilisation d'un correcticiel lors d'une épreuve de français au secondaire	6
Deuxième chapitre - Cadre conceptuel et état de la question.....	8
2.1 Correction par un humain.....	8
2.1.1 Caractéristiques des correcteurs.....	8
2.1.2 Processus de correction.....	9
2.1.3 Le concept d'erreur	11
2.1.4 Classification des erreurs	13
2.1.5 Exclusion de certaines erreurs.....	16
2.1.6 Erreurs courantes.....	16

2.2 Correction automatisée.....	17
2.2.1 Définition	17
2.2.2 Types de correcticiel	18
2.2.3 Fonctions des correcticiels	24
2.2.4 Processus de correction	25
2.2.5 Concept d'erreur dans le cadre d'un correcticiel	26
2.2.6 Classification des erreurs	26
2.2.7 Paramétrages de la correction	27
2.2.8 Avantages et limites	27
2.3 Comparaison entre deux correcticiels francophones.....	28
2.3.1 Correcteur 101.....	29
2.3.2 Antidote.....	31
2.4 Bilan et question de recherche	33
Troisième chapitre - Méthodologie de la recherche.....	35
3.1 Type de recherche	35
3.2 Élaboration du corpus de données.....	35
3.3 Description du fonctionnement du correcticiel.....	37
3.4 Description des analyses	42
3.5 Limites de la recherche	43
Quatrième chapitre - Résultats des analyses.....	45
4.1 Analyse quantitative.....	45
4.2 Analyse qualitative.....	57
4.2.1 La correction d'Antidote	57
4.2.2 La correction par les correcteurs humains	61
4.2.3 Retour sur les erreurs courantes	61
4.2.4 Différences dans le classement des erreurs selon le type de correcteur.....	63
Cinquième chapitre - Interprétation et discussion.....	66

Sixième chapitre - Conclusion.....	71
Bibliographie	xiii
Annexe 1: Grille de correction du MELS.....	xvi
Annexe 2: Grille pour l'analyse des erreurs.....	xix
Annexe 3: Autorisation de <i>Druide informatique inc</i>.....	xxi

Liste des tableaux

Tableau I	Pourcentage d'accord humain/ <i>E-rater</i> et humain.....	20
Tableau II	Performances du <i>Correcteur 101</i>	30
Tableau III	Fausses alertes commises par <i>Correcteur 101</i>	31
Tableau IV	Analyse des textes littéraires par les correcticiels.....	32
Tableau V	Analyse d'un article de journal par les correcticiels.....	33
Tableau VI	Analyse des compositions d'élèves par les correcticiels.....	33
Tableau VII	Correspondance entre les catégories.....	46
Tableau VIII	Nombre d'erreurs détectées selon les catégories.....	48
Tableau IX	Nombre d'erreurs détectées selon le niveau du texte.....	51
Tableau X	Nombre d'erreurs détectées selon les catégories – textes faibles.....	53
Tableau XI	Nombre d'erreurs détectées selon les catégories – textes moyens.....	54
Tableau XII	Nombre d'erreurs détectées selon les catégories – textes forts.....	55
Tableau XIII	Comparaison du niveau de performance attribué par les correcteurs.....	56
Tableau XIV	Observation des erreurs courantes dans les textes.....	62

Liste des figures

Figure 1	Théorie du processus cognitif chez le lecteur lors de la correction d'un texte.....	10
Figure 2	Écran de correction d'Antidote.....	39
Figure 3	Statistiques de correction – onglet Erreurs.....	40
Figure 4	Statistiques de correction – onglet Contenu.....	41
Figure 5	Nombre d'erreurs détectées selon les catégories.....	49
Figure 6	Nombre d'erreurs détectées selon le niveau du texte.....	51
Figure 7	Nombre d'erreurs détectées selon les catégories – textes faibles.....	53
Figure 8	Nombre d'erreurs détectées selon les catégories – textes moyens.....	54
Figure 9	Nombre d'erreurs détectées selon les catégories – textes forts.....	55

Liste des sigles et abréviations

Fig.	Figure
MELS	Ministère de l'Éducation, du Loisir et du Sport
TIC	Technologies de l'information et de la communication
*	L'astérisque précède une forme ou une expression fautive, une impropriété

Notation des erreurs détectées par les correcteurs humains (MELS)

G	Orthographe grammaticale
NC	Non classée
P	Ponctuation
S	Syntaxe
U	Orthographe d'usage
V	Vocabulaire

Notation des erreurs détectées par le correcticiel (*Antidote*)

Antidote-G	Orthographe grammaticale
Antidote-L	Lexique
Antidote-O	Orthographe d'usage
Antidote-P	Ponctuation
Antidote-R	Rédaction
Antidote-S	Syntaxe
AP	Analyse partielle
Style-Vabsent	Style (absence du verbe)
T-Casse	Typographie (casse)

*À Ariane
Pour sa présence tout au long
de la rédaction de ce mémoire*

Remerciements

Je tiens à remercier, tout d'abord, M. Jean-Guy Blais, mon directeur de recherche, de m'avoir proposé le sujet de ce travail. Je lui suis profondément reconnaissante pour son soutien et ses encouragements. Il a su me transmettre son intérêt pour la recherche en évaluation et je me sens privilégiée d'avoir pu bénéficier de ses précieux conseils.

Je souhaite également exprimer ma gratitude envers ma codirectrice, Mme Marie-Claude Boivin pour ses nombreuses relectures et ses commentaires des plus encourageants. Nos rencontres ont toujours été une source importante de stimulation dans mon travail.

Merci à l'équipe du GRIÉMÉtic pour la transcription des textes à l'ordinateur. Votre collaboration m'a permis de sauver de très nombreuses heures de travail.

Ce mémoire n'aurait pu se réaliser sans la collaboration du ministère de l'Éducation, du Loisir et du Sport qui nous a permis d'utiliser leurs données. Merci d'avoir rendu cette recherche possible.

Enfin, mille mercis à ma collègue et amie, Caroline Godin. Sans elle, je n'aurais osé mettre les pieds aux études supérieures. Sa présence à mes côtés pendant mes études a rendu le travail beaucoup plus agréable.

Introduction

L'évaluation des apprentissages est un domaine au cœur duquel de nombreux changements se produisent, et ce, depuis plusieurs années. La réforme scolaire n'est pas l'unique responsable de ce changement : l'évaluation doit aussi s'adapter aux changements qui se produisent dans la société. Ainsi, dans un monde où la technologie est de plus en plus mise au premier plan, le domaine de l'éducation doit inévitablement prendre un nouveau courant. Les élèves de demain auront peut-être tous un ordinateur portable dans les classes et il est possible qu'un jour, ils soient évalués à l'aide de divers programmes informatiques. Il faut donc envisager la conception de nouveaux outils d'évaluation davantage adaptés aux besoins futurs des enseignants.

Bien que l'utilisation des technologies soit présente dans de nombreuses disciplines scolaires, nous avons choisi de nous concentrer sur le français et l'évaluation des productions écrites dans cette langue. Nous avons découvert que plusieurs outils de correction automatique sont déjà utilisés, dans certains États américains, pour corriger à la fois le sens du texte et la qualité de la langue écrite. Notre objectif n'est pas, pour le moment, d'appliquer de tels outils lors des épreuves de français, mais plutôt d'étudier les différences entre un correcticiel et un correcteur humain quant à leur correction respective de la qualité de la langue de textes d'élèves du secondaire.

Les chapitres qui suivent présenteront les différentes étapes de notre recherche. Dans le premier chapitre, la problématique de la recherche sera abordée. Il sera entre autres question de l'utilisation de plus en plus présente des ordinateurs dans les écoles, de l'usage des correcticiels aux États-Unis, des avantages de ces correcticiels et des avantages possibles de l'utilisation d'un correcticiel lors d'une épreuve de français au secondaire. Le deuxième chapitre, formant le cadre conceptuel de cette recherche et

l'état de la question, sera divisé en deux parties principales, soit la correction par un humain et la correction automatisée. Nous aborderons ainsi, dans un premier temps, les caractéristiques des correcteurs humains, leur processus de correction, le concept d'erreur et la classification des erreurs qu'ils utilisent, l'exclusion de certaines erreurs lors de la correction des épreuves du MELS et les erreurs courantes retrouvées dans les textes d'élèves du secondaire. La seconde section traitera de la correction automatisée : nous définirons le concept de correcticiel et nous présenterons les différents types de correcticiel. Nous expliquerons ensuite leurs fonctions et leur processus de correction, le concept d'erreur dans le cadre d'un correcticiel, la façon dont les correcticiels, en général, classent les erreurs et les divers paramétrages de correction possibles. Nous dresserons également un portrait des avantages et des limites des correcticiels. Enfin, une troisième partie du cadre conceptuel présentera une comparaison entre les deux correcticiels francophones principalement utilisés au Québec, soit *Correcteur 101* et *Antidote*. Dans le troisième chapitre de ce mémoire, nous aborderons en détail la méthodologie en traitant de l'élaboration du corpus de données, du fonctionnement du correcticiel retenu (*Antidote*), des analyses effectuées et des limites de la recherche. Le quatrième chapitre présentera les résultats des analyses quantitatives et qualitatives. Enfin, le cinquième chapitre sera constitué de l'interprétation et de la discussion des résultats.

Premier chapitre

Problématique de la recherche

1.1 Utilisation des ordinateurs dans les écoles

Depuis plusieurs années, il est possible d'observer une utilisation accrue des ordinateurs dans les écoles. Cette situation est fortement encouragée par le ministère de l'Éducation, du Loisir et du Sport (MELS) qui, dans son Programme de formation de l'école québécoise, place l'exploitation des technologies de l'information et de la communication (TIC) parmi les compétences transversales que doivent développer les élèves. Selon le MELS, « le développement rapide des savoirs et l'accessibilité croissante à des sources d'information variées et multiples rendent l'utilisation des TIC de plus en plus incontournable » (MELS, 2006, p. 46). Puisque, comme le souligne le MELS, presque tous les emplois, à plus ou moins long terme, exigeront que les travailleurs utilisent au moins minimalement les technologies de l'information et de la communication, il est nécessaire de préparer les élèves à cette réalité. Ceci explique la place importante qu'occupent maintenant les ordinateurs dans les écoles. Les données du Conseil des ministres de l'Éducation indiquent qu'en 2003, un peu plus de 50 % des écoles utilisaient un ordinateur dans les cours de langue (Conseil des ministres, 2003).

Les élèves québécois apprennent à utiliser les ordinateurs pour leurs recherches et leurs travaux, mais ils continuent d'être évalués à l'aide d'un examen traditionnel présenté selon le format « papier et crayon »¹. Selon l'Organisation de coopération et de développement économique (OCDE), « dans la mesure où l'influence des TIC s'exerce dans tous les aspects de l'enseignement et de l'apprentissage et qu'elles

¹ L'expression « format papier et crayon » renvoie aux examens traditionnels pour lesquels les questions sont imprimées sur une feuille de papier. Les élèves doivent répondre aux questions sur cette même feuille ou sur une autre feuille réponse en utilisant un crayon, d'où l'appellation « papier et crayon ». Nous utiliserons cette expression pour parler de cette forme d'examen traditionnel.

deviennent un *support* de travail omniprésent, leur importance doit être proportionnellement prise en compte dans les procédures d'évaluation » (OCDE, 2001, p. 30). Il serait donc pertinent pour les élèves, qui utilisent de plus en plus les ordinateurs, de passer leurs examens sur ce support. Ainsi, en français, pour un examen de production écrite, les élèves pourraient rédiger leur texte à l'aide d'un logiciel de traitement de texte dans un laboratoire informatique ou dans une salle de classe, si des ordinateurs portables sont utilisés. Il pourrait être possible de bloquer l'utilisation des outils d'aide à la correction si l'on veut que les élèves utilisent uniquement leurs ouvrages de référence, comme c'est le cas dans les examens dits traditionnels. L'enseignant aurait ensuite le choix de corriger les textes imprimés ou directement à l'écran; dans ce cas, la possibilité d'utiliser un outil d'aide à la correction se présenterait à lui. Cette situation ne semble pas se vivre actuellement au Québec. Toutefois, ailleurs au Canada et aux États-Unis, certaines écoles ont déjà fait le saut et utilisent les technologies de l'information et de la communication lors des évaluations, que ce soit pour la production ou la correction de diverses formes d'épreuves.

1.2 Utilisation des correcticiels aux États-Unis

Depuis plusieurs années, les ordinateurs sont utilisés aux États-Unis en éducation, autant comme médium à partir duquel les élèves font leurs examens que comme outil de correction, lorsque des algorithmes spécialisés y sont installés. En effet, selon Bennett (2006), dans plusieurs États américains, les enseignants ont recours à certains logiciels nécessitant l'utilisation de l'ordinateur lors des évaluations. En 2006, 12 États avaient commencé à développer ou à administrer certaines composantes de leur programme d'évaluation officiel par le biais d'Internet, tandis que 8 autres États avaient mis en place des projets pilotes d'évaluation en ligne. Dans différents États américains, on peut constater l'implantation des épreuves administrées par le biais de l'ordinateur autant dans les écoles primaires que

secondaires, et ce, pour plusieurs matières, par exemple les mathématiques, les sciences, l'anglais et les sciences sociales.

Différents outils de correction automatique, appelés des correcticiels, ont aussi fait leur apparition. Parmi ceux-ci, les correcteurs de textes, en anglais, ont retenu l'attention de plusieurs chercheurs. Des outils de correction comme *PEG*, *IntelliMetric*, *E-rater* et *Intelligent Essay Assessor* ont fait l'objet de nombreuses études, et leur capacité à corriger la qualité de la langue et celle de l'argumentation dans des rédactions aussi efficacement que des correcteurs humains a été démontrée² (Dikli, 2006).

1.3 Avantages des correcticiels

Si les correcticiels sont aussi populaires aux États-Unis, c'est parce qu'ils offrent de nombreux avantages. Ils permettent tout d'abord une économie de temps et d'argent (Dikli, 2006). En effet, un texte qui prendrait une vingtaine de minutes à évaluer par un correcteur humain peut être corrigé en quelques secondes par un correcticiel. Ainsi, lors des épreuves officielles qui demandent l'utilisation de deux correcteurs pour chaque examen, le remplacement d'un des deux correcteurs par un outil informatique pourrait être une solution très économique. Les correcticiels ont aussi l'avantage d'être plus constants que les correcteurs humains. Les correcticiels ne sont pas influencés par la fatigue, les délais à respecter ou tout autre biais : ils peuvent donc évaluer les textes de façon plus constante et objective (Foltz, Laham et Landauer, 1999). La fiabilité des correcteurs humains est beaucoup moins sûre, puisque, contrairement aux correcticiels, ils peuvent être sujets à la fatigue et à l'effet halo. Selon Page, l'effet halo consiste en une tendance qu'ont les correcteurs humains à établir une corrélation élevée entre les différents traits qu'ils évaluent, soit la cohérence du texte, la structure, la ponctuation, la syntaxe, etc. (Page, 1994). Ainsi,

² La section 2.2.2 de ce mémoire présentera en détail le fonctionnement de ces correcticiels et les études qui ont été réalisées pour prouver leur efficacité.

un correcteur humain aura tendance à mettre des notes trop élevées à certains critères si l'ensemble des critères est évalué de façon élevée et, inversement, il aura tendance à mettre des notes trop faibles à certains critères si l'ensemble des autres critères est faible. La situation inverse peut également être observée à travers un phénomène de « compensation » lorsqu'un même correcteur a la responsabilité de l'ensemble des critères de correction. Le correcteur qui trouve que le candidat obtient des résultats faibles pour certains critères appréciés au départ voudra donner une chance au candidat et pourra être plus généreux pour les critères appréciés en dernier. Les correcticiels, contrairement aux humains, ne portent pas de jugement et ne peuvent donc pas être soumis à ce type d'effet. Leur constance est ainsi assurée.

1.4 Avantages possibles de l'utilisation d'un correcticiel lors d'une épreuve de français au secondaire

Au Québec, à notre connaissance, on n'utilise pas encore, du moins officiellement, la correction automatique des textes. Pourtant, certains examens pourraient bénéficier, tout comme aux États-Unis, des avantages des correcticiels. La situation qui nous intéresse concerne une épreuve uniforme de français au secondaire soumise par le ministère de l'Éducation, du Loisir et du Sport afin de proposer un prototype de bilan de fin de cycle. Avec la réforme scolaire actuelle, le ministère a voulu mettre à la disposition des enseignants des prototypes d'épreuves facultatifs permettant l'évaluation des élèves à la fin des cycles. L'épreuve prototypique utilisée pour notre recherche a pour but de juger de la compétence à écrire un texte descriptif pour des élèves de deuxième secondaire. Il s'agit donc d'un outil précieux pour faire le point sur les difficultés d'écriture des élèves. Cette épreuve est accompagnée d'un guide de correction pour les enseignants et les correcteurs du ministère indiquant les différents éléments dont il faut tenir compte lors de la correction (MELS, 2008). Une grille de correction³ est également fournie afin de compiler les résultats des élèves dans les différents volets de l'écriture. Le guide de correction indique que la

³ Voir *Annexe I*

correction des erreurs de langue est faite avec précision et que chaque erreur est classée selon une catégorie, soit *syntaxe*, *ponctuation*, *orthographe grammaticale* ou *d'usage*, etc. Toutefois, la grille de correction n'est pas aussi détaillée et ne prévoit que la comptabilisation des erreurs selon les catégories *construction de phrases* et *respect des normes relatives*. Il est donc impossible de savoir, à partir de cette grille, avec quels aspects de la langue les élèves ont le plus de difficultés. Un outil de correction automatique pourrait offrir l'avantage d'une analyse plus constante et plus détaillée des erreurs, ce qui permettrait ensuite aux enseignants de mettre davantage l'accent sur certaines notions, et ce, beaucoup plus tôt dans le cursus scolaire, afin d'aider les élèves. Un correcticiel qui ferait la correction de la qualité de la langue, utilisé conjointement avec les correcteurs humains, pourrait également permettre à ceux-ci de se concentrer uniquement sur le contenu en laissant le soin au correcticiel de corriger les erreurs de langue. Il y aurait ainsi une meilleure utilisation des ressources disponibles, soit une économie de temps et, par le fait même, une économie d'argent. Il faut toutefois se demander si les correcticiels francophones actuellement disponibles sur le marché sont assez développés pour pouvoir corriger efficacement les erreurs de langue dans des textes d'élèves du secondaire. Les avantages que ces correcticiels peuvent offrir sauraient-ils contrebalancer leurs éventuelles limites? Si ce n'est pas le cas, les coûts associés à l'implantation d'un outil de correction plus sophistiqué en vaudraient-ils la peine?

En somme, il nous semble pertinent d'appliquer un correcticiel sur les textes d'élèves lors d'une épreuve de français et d'en étudier les résultats. Ces résultats seront mis en parallèle avec ceux obtenus lors de la correction habituelle effectuée par un correcteur humain afin d'établir l'existence d'éventuelles différences et de juger de l'utilité de cet outil de correction lors des épreuves de français. La question qui nous préoccupe est donc la suivante : quelles sont les différences entre la correction humaine et la correction effectuée par un correcticiel francophone lors de la correction de la qualité de la langue dans des textes d'élèves du secondaire?

Deuxième chapitre

Cadre conceptuel et état de la question

Afin d'étudier les résultats obtenus lors de l'application d'un correcticiel sur une épreuve uniforme de français au secondaire, il est nécessaire de bien comprendre le processus de correction des correcteurs humains et de le comparer à celui des correcticiels.

2.1 Correction par un humain

Dans cette section, les différents concepts permettant de comprendre la correction effectuée par des correcteurs humains seront présentés. Les caractéristiques des correcteurs seront tout d'abord décrites, suivies par le processus de correction humaine. Le concept d'erreur pour un correcteur humain sera ensuite abordé, de même que la classification des erreurs et les erreurs courantes retrouvées dans les textes d'élèves.

2.1.1 Caractéristiques des correcteurs

L'expression « correcteur humain », telle que rencontrée dans la littérature⁴, peut désigner diverses personnes de formation et de compétence différentes. Il peut s'agir d'enseignants du primaire ou du secondaire, en français ou dans une autre matière, de professeurs au Cégep ou à l'université. Il peut également s'agir de linguistes ou de traducteurs. Dans le cas qui nous intéresse, les correcteurs humains

⁴ Voir, entre autres, Jacquet-Pfau (2001) et Dessus et Lemaire (2004).

seront définis comme des professionnels ayant reçu une formation spécifique pour corriger les épreuves de français au secondaire.

2.1.2 *Processus de correction*

Le processus de correction, pour un correcteur humain, est un processus complexe qui fait appel à plusieurs activités cognitives à la fois. Lors de la correction d'un texte, un correcteur humain doit porter son attention sur divers éléments de contenu et de forme. Les éléments de contenu sont davantage liés à une idée d'excellence à atteindre, tandis que les éléments de forme renvoient à une structure de la langue précise qu'il faut respecter. Même si, pour cette recherche, nous désirons nous attarder à la correction de la forme du texte, nous devons aborder le processus de correction dans son ensemble, puisque les correcteurs humains qui doivent corriger des textes d'élèves du secondaire effectuent une tâche de correction complète en évaluant à la fois le contenu et la forme du texte.

Selon Cordier-Gauthier et Dion (2003), « les paramètres de la correction humaine sont les mêmes au départ que ceux impliqués dans le processus de la lecture ». Cependant, l'intention de lecture du correcteur est double, puisqu'il tente à la fois de comprendre le message et de l'évaluer. La complexité de la correction humaine vient aussi du fait que le correcteur doit sans cesse jouer un double rôle. Il doit jouer le rôle « d'intervenant au plan de la norme linguistique » (Cordier-Gauthier et Dion, 2003) en proposant des solutions ou des corrections à la fois justes et adaptées et il doit également jouer le rôle d'intervenant pédagogique en tenant compte du niveau attendu de l'élève. La correction humaine est donc considérée comme complexe à cause des deux activités cognitives qui se déroulent en même temps, soit la compréhension du message et le jugement, ainsi qu'à cause des deux rôles que doit jouer le correcteur, soit celui de faire respecter la norme linguistique et celui de tenir compte du niveau attendu de l'élève.

La complexité de la correction humaine peut également être expliquée par l'existence possible d'une différence entre les diverses représentations mentales que se font les correcteurs face à un même texte. Wolfe (1997) parle des deux activités cognitives qui entrent en interaction en utilisant les termes « framework of writing » et « framework of scoring ». Le « framework of writing » se définit comme la représentation mentale que se fait le correcteur par rapport aux critères de correction. Ces critères regroupent des éléments positifs, comme l'apparence du texte, l'organisation, la structure, l'orthographe, etc., et d'autres négatifs, comme l'imprécision. Le « framework of scoring », quant à lui, est une représentation créée par le correcteur par rapport au processus mental utilisé pour corriger un texte. Il regroupe l'interprétation, l'évaluation et la justification. La figure ci-dessous (*Fig. 1*) illustre la théorie de Wolfe.

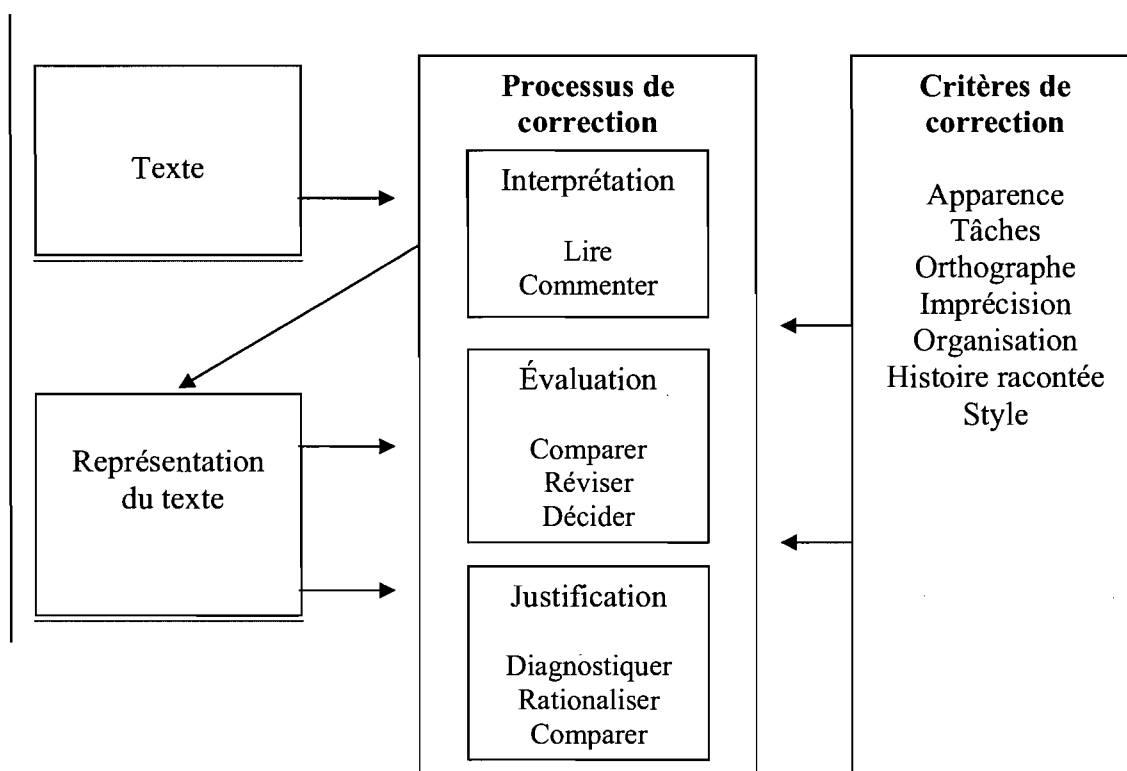


Fig. 1 - *Théorie du processus cognitif chez le lecteur lors de la correction d'un texte**

* Traduction de l'auteure

Source : Wolfe, E. W. (1997). The relationship between reading style and scoring proficiency in a psychometric scoring system. *Assessing Writing*, 4, p. 89.

Le texte que reçoit le correcteur peut donc, selon cette théorie, être interprété différemment selon les représentations mentales que le correcteur se crée en ce qui a trait aux critères et au processus de correction. Les critères de correction n'étant pas toujours uniformes entre les correcteurs, le « framework of writing » peut varier d'un correcteur à l'autre, ce qui peut amener une interprétation différente du texte et, par le fait même, une évaluation différente.

À ceci doivent également s'ajouter les différents facteurs externes qui peuvent venir influencer le correcteur, soit les facteurs psychologiques comme la lassitude, le manque de tolérance ou l'ennui et les facteurs physiologiques comme la fatigue. Ces facteurs, ajoutés à la complexité du processus de correction et aux différentes interprétations possibles d'un texte, expliquent pourquoi la correction humaine n'est pas infaillible et pourquoi deux correcteurs n'ont pas toujours le même regard sur un texte.

2.1.3 *Le concept d'erreur*

Pour définir le concept d'*erreur*⁵, il faut inévitablement se tourner vers les grammaires. Ces dernières peuvent être descriptives ou normatives. Selon Riegel (2004), une grammaire descriptive rend compte des régularités de la langue. Son point de vue est, tel que son nom l'indique, purement descriptif. Il en va toutefois autrement avec les grammaires normatives, qui se donnent comme rôle de refléter (et d'enseigner) le bon usage de la langue. Le concept d'*erreur* est donc inévitablement relié à celui de *norme*. La grammaire normative est alimentée par la norme prescriptive (Anctil, 2005), norme qui dicte le bon usage. Les variantes linguistiques qui ne correspondent pas à l'usage prescrit sont condamnées et elles sont

⁵ Bien que le terme « faute » soit souvent utilisé au même titre que le terme « erreur », nous choisissons, pour cette recherche, de parler d'erreur puisque ce terme est, en général, beaucoup moins connoté péjorativement. En effet, en didactique, le terme d'erreur a un sens dans l'apprentissage, révélant une construction partielle des connaissances. Les erreurs font partie du processus normal d'apprentissage. Ainsi, puisque notre recherche touche à des textes d'élèves du secondaire étant encore en apprentissage de la langue, le terme d'erreur s'applique bien.

automatiquement qualifiées d'*erreurs*. Une grammaire normative va donc être beaucoup plus sévère et qualifier d'erreur ce qui pourrait être tout à fait adéquat dans certaines situations. Soulignons que la norme dite *descriptive* est fort différente de la norme prescriptive. Selon cette norme, une erreur correspond non pas à ce qui s'écarte du bon usage, mais plutôt à ce qui s'écarte de la majorité (Anctil, 2005). Ainsi, si la majorité de la population, dans un milieu donné, écrit **une avion*, cette forme deviendrait la norme. Notre présente recherche concernant des correcteurs du MELS, nous devons nous attendre à ce que la norme prescriptive, dictée dans des grammaires normatives, soit utilisée. En effet, le système scolaire a pour rôle de faire apprendre la norme prescriptive aux élèves : on n'enseignerait jamais **une avion*, tant que la norme prescrit *un avion*, puisqu'on veut transmettre cette norme. C'est pourquoi nous nous contenterons de développer le concept d'erreur selon des grammaires normatives, comme celle de Riegel.

Selon Riegel, les « vraies fautes contre la langue » sont des « formes irrécourables qui contreviennent aux règles communes à l'ensemble des sous-systèmes d'une même langue » (Riegel, 2004, p. 15). Les erreurs peuvent également concerner non pas les formes, mais plutôt leur emploi inadéquat. Il y a d'ailleurs une différence importante entre des formes considérées comme inacceptables et des formes considérées comme agrammaticales. Riegel donne l'exemple suivant pour illustrer cette différence :

L'élève dont le devoir que j'ai lu hier soir était mauvais est votre fils.

(Riegel, 2004, p. 19)

Selon Riegel, cette phrase est grammaticale, mais inacceptable. En effet, c'est la complexité de la structure qui rend la phrase inacceptable, puisque la grammaticalité est bien respectée. Une phrase acceptable, selon Riegel, est « une phrase pour laquelle il n'y aurait aucune difficulté à imaginer un ou des contextes où son interprétation ne poserait pas de problème. » (Riegel, 2004, pp. 19-20) La grammaticalité, quant à elle, concernerait « les règles de bonne formation intrinsèque des énoncés » (Riegel, 2004, p. 20), soit les règles morphologiques, syntaxiques, sémantiques et, selon

certaines grammaires, pragmatiques. L'exemple suivant illustre une phrase agrammaticale :

**La hausse des prix des aliments ont surpris les consommateurs.*

Cette phrase est considérée agrammaticale puisqu'elle présente une erreur dans l'accord du verbe. En effet, le verbe *surprendre* devrait s'accorder avec *hausse*, 3^e personne du singulier. Les règles de formation des énoncés ne sont pas respectées, d'où l'agrammaticalité de la phrase.

Le concept d'erreur, pour un correcteur humain, se rattache donc, d'une part, à ce qui ne correspond pas aux règles de la langue (ce qui est agrammatical) et, d'autre part, à ce qui est soit inacceptable, soit inapproprié dans le contexte de communication.

2.1.4 *Classification des erreurs*

En plus de devoir identifier correctement les erreurs dans un texte, les correcteurs humains doivent également les classer selon certaines catégories. Cette classification varie parfois d'une grammaire à l'autre et également d'un correcteur à l'autre. Nous examinerons, dans cette section, les différences entre la classification faite par une grammaire de référence, soit celle de Riegel, une grammaire descriptive, et la classification faite par le ministère de l'Éducation, du Loisir et du Sport dans le guide de correction de l'épreuve d'écriture proposée aux élèves du 1^{er} cycle du secondaire.

Riegel définit tout d'abord la syntaxe comme « la façon dont les mots se combinent pour former des groupes de mots et des phrases » (Riegel, 2004, p. 22). La ponctuation, quant à elle, est définie comme « le système des signes graphiques qui contribuent à l'organisation d'un texte écrit et qui apportent des indications prosodiques, marquent des rapports syntaxiques ou véhiculent des informations

sémantiques » (Riegel, 2004, p. 84). Dans la grille de correction de l'épreuve que propose le MELS, la construction des phrases et la ponctuation sont regroupées ensemble sous un même critère. Le MELS utilise la définition du Petit Robert pour définir le concept de syntaxe : « Étude des règles qui président à l'ordre des mots et à la construction des phrases » (MELS, 2008, p. 14). Pour le MELS, les phrases sont considérées « correctes » si les critères suivants sont respectés :

- « - Tous les mots nécessaires doivent être présents.
- L'ordre des mots est correct.
- Les auxiliaires, les verbes transitifs directs et indirects et les verbes attributifs sont bien employés.
- Les temps verbaux sont bien harmonisés.
- Les pronoms, les prépositions et les conjonctions sont bien employés.
- Les pronoms et les déterminants à la troisième personne rappellent clairement des êtres ou des idées déjà mentionnés dans le texte.
- Le mode, la voix et le temps des verbes sont bien employés. » (MELS, 2008, p. 14)

Le MELS fournit ensuite aux correcteurs une série de principes à respecter, comme celui de ne compter qu'une erreur de syntaxe par phrase syntaxique. Bien que le MELS soit plus précis que Riegel quant aux types d'erreurs de syntaxe, certains types d'erreurs qu'il inclut dans sa définition pourraient être classés ailleurs. Par exemple, les auxiliaires, les verbes transitifs et attributifs, les pronoms, les prépositions et les conjonctions pourraient être classés en lexicale, tandis que les temps verbaux pourraient être classés en cohérence.

En ce qui concerne la ponctuation, le MELS ne donne pas de définition, mais il indique tous les cas où la virgule doit être obligatoire et ceux où elle est facultative. Il fournit également une liste de cas à exclure. Les correcteurs ne doivent donc pas corriger certains cas plus difficiles en ponctuation, comme l'emploi du point-virgule.

L'orthographe d'usage et l'orthographe grammaticale ne sont pas définies dans la grammaire de Riegel. De Villers (2003) fournit toutefois une définition assez précise. Elle explique que l'orthographe d'usage, également nommée orthographe usuelle ou orthographe lexicale, « concerne la graphie d'un mot indépendamment du rôle qu'il joue dans la phrase » (De Villers, 2003, p. 1030). L'orthographe grammaticale, quant à elle, « régit la graphie d'un mot selon le rôle qu'il joue dans la phrase et définit donc les accords » (De Villers, 2003, p. 1030). Le MELS regroupe sous la rubrique « orthographe » l'orthographe d'usage et l'orthographe grammaticale. On s'attend à ce que l'élève soit capable d'orthographier correctement les termes courants. Selon le MELS, les termes courants réfèrent aux « mots qui sont communs, usuels, habituels et dont l'occurrence est élevée dans la langue » (MELS, 2008, p. 20) ainsi que les « mots qui peuvent provenir de champs lexicaux liés à certains thèmes » (MELS, 2008, p. 20). L'élève doit également effectuer les accords usuels dans le groupe du nom et le groupe du verbe, c'est-à-dire les accords en genre et en nombre des noms⁶, des adjectifs, des participes passés employés seuls, avec l'auxiliaire avoir et avec l'auxiliaire être. Il doit aussi effectuer correctement les « accords des verbes réguliers et irréguliers usuels aux temps les plus courants » (MELS, 2008, p. 20) et utiliser la terminaison verbale appropriée.

En somme, la classification des erreurs est semblable entre les deux grammaires consultées et la grille de correction du MELS. Toutefois, dans la grille du MELS, certaines erreurs sont exclues afin de ne pas pénaliser les élèves sur des notions qu'ils n'ont pas encore apprises.

⁶ Soulignons que les accords des noms dont le MELS fait mention réfèrent plutôt à l'accord des déterminants et à la présence des marques morphologiques de nombre (et de genre, le cas échéant) sur les noms, puisque les noms sont des donneurs d'accord (ils ne « s'accordent » donc pas, c'est-à-dire qu'ils ne reçoivent pas d'accord).

2.1.5 Exclusion de certaines erreurs

Dans son guide de correction, le MELS fournit une liste d'erreurs à ne pas prendre en compte dans les textes. Nous retrouvons dans cette liste, entre autres, le pluriel des noms composés, le pluriel des adjectifs composés de couleur, le participe passé des verbes pronominaux, etc.. (MELS, 2008, p. 22). Cette liste est établie en fonction des attentes propres au niveau scolaire des élèves. Le correcteur qui voit ces erreurs peut les identifier, mais il ne les comptabilise pas. Soulignons que cette possibilité d'identifier une erreur sans la comptabiliser est propre aux correcteurs humains, les correcticiels ne pouvant décider de comptabiliser ou non les erreurs.

2.1.6 Erreurs courantes

Puisque notre recherche porte principalement sur les textes d'élèves du secondaire, il est intéressant de connaître les erreurs les plus courantes sur lesquelles nous devrions porter notre attention. Boivin, Pinsonneault et Philippe (2003) ont relevé certaines erreurs courantes. Les exemples suivants sont tirés de leur ouvrage intitulé *Bien écrire, la grammaire revue au fil des textes littéraires* (les phrases agrammaticales sont précédées du symbole *). Dans le domaine de la syntaxe de la phrase, plusieurs erreurs touchent les constructions des phrases négatives (cf. **J'ai pas vu personne dans le parc* vs *Je n'ai vu personne dans le parc*) et les constructions des phrases interrogatives (cf. **Ma réaction le surprend-il?* vs *Ma réaction le surprend-elle?*). Dans le groupe nominal, les erreurs courantes concernent le genre du nom, comme dans les exemples : **une autobus*, **une avion*, **une bel horaire*, etc. Les phrases complexes, quant à elles, présentent souvent des erreurs dans le choix du mode dans la phrase subordonnée, comme dans la phrase **Cela ne sera pas beau, à moins qu'on met une photo*, dans la construction des complétives interrogatives indirectes, comme **Paul ignore où est-ce que Marie ira* ou dans le choix du pronom relatif, comme **Le livre que j'ai besoin a été volé*. Puisque tous ces éléments

représentent des sources d'erreur courantes, il sera important de vérifier si les systèmes de correction automatique sont capables de les repérer.

Dans leur ouvrage intitulé *La grammaire nouvelle - La comprendre et l'enseigner*, Nadeau et Fisher (2006) traitent, elles aussi, de certaines difficultés de la langue rencontrées fréquemment dans les textes d'élèves. Elles parlent, entre autres, du problème concernant les accords de proximité. Ce problème survient lorsque le sujet du verbe est un groupe nominal complexe, comme dans l'exemple suivant : *La hausse des prix des aliments limite le choix des consommateurs*. Il est fréquent de retrouver l'erreur suivante : **La hausse des prix des aliments limitent le choix des consommateurs*. En effet, les noms pluriels *prix* et *aliments* se trouvant plus près du verbe, plusieurs vont faire un accord erroné.

Il ne s'agit bien évidemment pas de la liste complète des erreurs commises par des élèves du secondaire, mais cette liste nous donne un aperçu des erreurs possibles tout en permettant de nous orienter pour l'analyse de l'efficacité du correcticiel que nous choisirons d'utiliser pour notre recherche.

2.2 Correction automatisée

2.2.1 Définition

Les systèmes de correction automatisée sont souvent désignés sous le nom de « correcticiel » ou de « correcteur ». Charnet et Panckhurst (1998) définissent le « correcteur grammatical » comme « un logiciel qui permet de vérifier et de corriger un texte en langue naturelle, saisi sur format électronique, par exemple, au sein d'un traitement de texte. » Cette définition touche à deux fonctions des correcticiels, soit, en premier lieu, la vérification, puis la correction proprement dite du texte. Il s'agit donc d'un logiciel qui ne fait pas que repérer les erreurs, mais qui peut également proposer une correction. Cette définition rejoint celle de Désilets (1999) qui utilise les

termes « outils d'aide à la correction », « correcteur » ou « correcticiel » pour parler de « logiciels capables de détecter les erreurs contenues dans un document rédigé à l'ordinateur et éventuellement de proposer des correctifs. » Selon Désilets, « ces outils peuvent avoir un fonctionnement autonome, mais ils sont de plus en plus souvent intégrés aux logiciels de traitement de texte à l'intérieur d'un menu. Leur domaine d'intervention peut concerner aussi bien l'orthographe lexicale et grammaticale que la syntaxe et même la stylistique (longueur et complexité des phrases, registres de langue, utilisation de la voix passive, etc.) » (Désilets, 1999, p. 3). La définition de Désilets rejoint donc celle de Charnet et Panckhurst en ce qui a trait aux deux principales fonctions du correcticiel, mais elle va un peu plus loin en mentionnant leur domaine d'intervention. Une autre définition, celle de Morin (1995), fait également état des domaines d'intervention du correcticiel. Selon Morin, « on entend par "correcticiel" tout logiciel de correction de la langue écrite capable d'intervenir non seulement sur l'orthographe des mots, mais aussi sur d'autres éléments comme la syntaxe, la ponctuation, voire même le sens » (Morin, 1995, p. iii). Morin va donc un peu plus loin dans sa définition puisqu'il parle des systèmes de correction qui sont capables d'évaluer le sens du texte. En effet, certains correcticiels sont en mesure de juger si le texte est cohérent, si les idées s'enchaînent correctement, si les informations sont pertinentes, etc. Cette forme de correcticiel est surtout utilisée dans la langue anglaise. Les outils de correction francophones sur lesquels nous désirons nous pencher se limitent à la correction de l'orthographe, de la syntaxe et de la stylistique. Ainsi, dans le cadre de notre recherche, nous privilégions la définition de Désilets et nous utiliserons le terme « correcticiel » pour parler des outils de correction.

2.2.2 Types de correcticiel

Comme nous l'avons mentionné précédemment, il existe une différence importante entre les correcticiels anglophones et francophones. Les correcticiels développés et utilisés principalement aux États-Unis sont conçus pour corriger à la

fois la forme et le sens des textes. À ce jour, les principaux systèmes de correction anglophones sont *Project Essay Grade (PEG)*, *E-rater*, *Intellimetric* et *Intelligent Essay Assessor*.

PEG est le premier système de ce genre à avoir été développé aux États-Unis. Selon Page (2003), le but de *PEG* est de prédire le résultat qu'un certain nombre de correcteurs compétents donneraient à un ensemble de textes similaires. Le système de correction ne peut pas véritablement analyser les variables intrinsèques d'un texte, c'est-à-dire les qualités se retrouvant dans l'écriture, comme la fluidité, l'organisation, etc. *PEG* utilise plutôt des approximations de ces variables, comme la longueur des mots, le nombre de points-virgules ou la rareté des mots (Rudner et Gagne, 2001). Le correcticiel se sert donc de ce qui peut être observable pour baser son analyse. Est-ce surprenant de constater que des critères comme la longueur des mots peuvent, à la suite de complexes combinaisons, prédire la note qu'un correcteur humain attribuerait au texte? Oui, sans aucun doute. *PEG* a toutefois su faire ses preuves et son efficacité a été démontrée. Shermis, Mzumara, Olson et Harrington (2001) ont étudié le fonctionnement et l'efficacité de *PEG* afin d'évaluer la possibilité d'utiliser ce correcticiel pour des examens réalisés sur ordinateur. Leur étude est formée de deux expériences. Dans la première, ils ont utilisé 1293 textes d'élèves du secondaire et du collégial qui avaient préalablement été corrigés par six correcteurs humains pour créer un modèle statistique qui allait être intégré au logiciel de correction. À partir de ces 1293 textes déjà corrigés, *PEG* a pu définir 30 variables observables sur lesquelles appuyer ses prochaines analyses. Dans la seconde expérience, 617 textes ont été soumis au correcticiel. Ces textes avaient également été corrigés par six correcteurs humains et les chercheurs ont voulu comparer les résultats à l'aide de tests statistiques. Cette étude a montré que les notes attribuées par *PEG* s'approchaient de la moyenne des jugements humains. Il y a en fait moins de différence entre le jugement de *PEG* et celui de la moyenne des correcteurs humains qu'entre les jugements des différents correcteurs humains entre eux.

Le second correcticiel, *E-rater*, est un système qui compare les différents éléments d'un texte avec ceux contenus dans un ensemble de textes de référence corrigés par des correcteurs humains, pour ensuite attribuer au texte une note de 1 à 6 (Burstein, Kukich, Wolff, Lu et Chodorow, 1998). Plus de 100 variables sont utilisées pour établir la note attribuée. Elles sont combinées entre elles au moyen d'une régression linéaire conduisant à un modèle de correction le plus près possible de celui des correcteurs humains (Kukich, 2000). L'efficacité de ce système de correction a été prouvée à maintes reprises, entre autres par Hutchison, en 2007. Ce dernier a voulu comparer les notes attribuées par des correcteurs humains et par *E-rater* à des textes d'élèves de 11 ans. Il a recueilli 600 textes rédigés à la main qu'il a soumis à trois correcteurs humains. Les textes devaient être notés de 1 à 5 à partir de critères tels la ponctuation, la syntaxe, le style d'écriture et l'organisation du texte. Les textes étaient ensuite retranscrits à l'ordinateur afin d'être corrigés par *E-rater*. Les résultats de l'étude de Hutchison sont regroupés dans le tableau I présenté ci-dessous.

Tableau I
Pourcentage d'accord humain/*E-rater* et humain/humain

Accord entre		Exact	Exact + adjacent (1 point de différence)
<i>E-rater</i>	Correcteur 1	40.46	87.46
<i>E-rater</i>	Correcteur 2	41.03	86.32
<i>E-rater</i>	Correcteur 3	40.46	87.46
<i>E-rater</i>	Moyenne des correcteurs	45.87	89.46
Correcteur 1	Correcteur 2	47.58	86.89
Correcteur 1	Correcteur 3	41.88	88.03
Correcteur 2	Correcteur 3	37.89	82.62

Source : Hutchison, D. (2007). An evaluation of computerized essay marking for national curriculum assessment in the UK for 11-years-old. *British Journal of Educational Technology*, 38(6), p. 982.

Ces résultats montrent que le pourcentage d'accord exact entre *E-rater* et la moyenne des correcteurs humains (45,87 %) est plus élevé que le pourcentage d'accord exact entre les correcteurs 1 et 3 (41,88 %) et les correcteurs 2 et 3 (37,89 %). Le pourcentage d'accord exact est toutefois plus élevé entre les correcteurs 1 et 2 (47,58 %) qu'entre *E-rater* et la moyenne des correcteurs. Autrement dit, la différence des résultats entre *E-rater* et la moyenne des correcteurs humains est moins élevée qu'entre deux correcteurs humains dans deux cas sur trois. *E-rater* a donc, dans cette étude, une efficacité comparable à celle des correcteurs humains. D'autres chercheurs ne sont toutefois pas aussi certains de l'efficacité de ce correcticiel. Powers, Burstein, Fowles, Chodorow et Kukich, en 2001, ont voulu tester les limites d'*E-rater*. Pour leur étude, ces chercheurs ont invité différents experts de l'écriture à tenter de piéger le correcticiel. Leur échantillon de textes contenait, entre autres, un essai composé de quelques paragraphes répétés 37 fois. Les correcteurs humains ont attribué une note de 1 à ce texte, tandis que le correcticiel lui a attribué une note de 6, soit le plus haut résultat possible. Les chercheurs ont donc conclu que, bien que l'efficacité d'*E-rater* ait été souvent démontrée, son emploi devrait se faire conjointement à celle d'un correcteur humain afin d'éviter que de telles situations se produisent. Les concepteurs d'*E-rater* auraient toutefois développé, depuis, de nombreux filtres qui permettent de signaler les textes douteux comportant des répétitions excessives (Powers et al., 2001).

Intellimetric est un système qui, comme *E-rater*, doit être entraîné à partir de textes corrigés par des correcteurs humains. Il utilise un mélange d'intelligence artificielle et de technologies statistiques (Elliot, 2003). Elliot a présenté plusieurs études réalisées par Vantage Learning, la compagnie ayant créé *Intellimetric*, qui ont montré l'efficacité du correcticiel. L'une d'entre elles utilisait une approche semblable à celle utilisée par Hutchison pour *E-rater*. Environ 612 textes rédigés dans le cadre d'un examen national K-12⁷ en 1998 ont été sélectionnés pour cette étude. Chaque texte était corrigé par un correcteur humain et une note de 1 à 6 lui était attribuée. Contrairement aux études présentées pour les autres correcticiels, cette étude de

⁷ Examens administrés dans différentes matières aux élèves de la 1^{re} à la 12^e année.

Vantage Learning comparait les résultats d'*Intellimetric* à ceux d'un seul correcteur humain. Sur les 612 textes retenus pour l'étude, 102 étaient utilisés pour programmer le correcticiel. Les résultats ont montré que le pourcentage d'accord exact entre *Intellimetric* et le correcteur humain variait entre 90 et 98 %. Certains chercheurs, comme Wang et Brown (2007) ont toutefois émis des réserves face aux études montrant l'efficacité des correcticiels, études réalisées en majorité par les concepteurs eux-mêmes. Elles ont donc fait une étude visant à comparer la note moyenne attribuée par *Intellimetric* et par 2 correcteurs humains à un groupe de textes rédigés par des élèves d'origine hispanique ou mexicaine ayant l'anglais comme langue première. Leur étude a entre autres démontré qu'*Intellimetric* attribuait une moyenne générale statistiquement plus élevée que celle attribuée par les correcteurs humains. Elles ont aussi constaté que, sur un échantillon de 107 textes évalués sur une échelle de 1 à 8, *Intellimetric* attribuait des résultats de 4 ou moins à 3 textes, tandis que les correcteurs humains donnaient des notes de 4 ou moins à 29 textes. Il est donc évident qu'*Intellimetric* était moins sévère, dans cette expérience, que les correcteurs humains. Wang et Brown attribuent ce résultat au fait que les textes ayant servi à calibrer le correcticiel ne proviennent pas de la même population que ceux utilisés dans leur recherche. Selon elles, *Intellimetric* effectue une correction plus juste lorsqu'il est calibré avec des textes provenant d'une même population que ceux qu'il évalue. Ainsi, si les textes ayant servi à calibrer le correcticiel ont été rédigés par un groupe d'élèves d'origine américaine fréquentant des collèges privés, il pourrait y avoir un biais lorsqu'on tenterait d'utiliser le correcticiel pour corriger des textes d'élèves provenant d'un autre type de population. C'est du moins l'hypothèse que ces chercheuses ont émise à la suite de leur recherche.

Enfin, le quatrième système de correction utilisé principalement aux États-Unis, *Intelligent Essay Assessor*, utilise un processus connu sous le nom de « Latent semantic analysis » (Landauer, Laham et Foltz, 2003). Nichols, en 2005, a mené une étude sur ce correcticiel. Il a utilisé un total de 3244 textes répartis selon trois niveaux scolaires. Ces textes étaient rédigés à la main par les élèves et retranscrits ensuite sur ordinateur afin que le correcticiel puisse les corriger. Les textes ont été corrigés par

des correcteurs humains experts, par des lecteurs ayant reçu un entraînement à la correction et par *Intelligent Essay Assessor*. L'étude regroupait 53 lecteurs et 2 correcteurs experts. Chaque texte rédigé à la main était tout d'abord corrigé de façon indépendante par chacun des deux correcteurs experts. Ceux-ci devaient attribuer une note de 1 à 6 pour chacun des textes. Nichols précise que les correcteurs experts étaient en parfait accord 96 % du temps. Chaque texte était ensuite corrigé par deux lecteurs, selon la même grille d'évaluation, puis enfin par le correcticiel. À la suite d'une série d'analyses statistiques, Nichols a pu montrer que le taux d'accord exact entre la note attribuée par le correcticiel et par le correcteur expert est de 59,62 %, tandis que le taux d'accord adjacent (un point de différence) est de 97,6 %. Le taux d'accord exact entre le correcticiel et le correcteur expert, même s'il est moins élevé que celui entre deux correcteurs experts, est tout de même plus élevé que celui entre un lecteur entraîné et un correcteur expert (50 %). Le correcticiel est donc plus efficace qu'un lecteur entraîné à la correction. Nichols explique la plus grande efficacité du logiciel par rapport aux lecteurs entraînés par le fait que le correcticiel a un processus de correction beaucoup plus semblable à celui des correcteurs experts que ne l'ont les lecteurs entraînés, mais peu expérimentés. Les correcteurs experts ont tendance à interpréter l'ensemble du texte, puis à l'évaluer. Le correcticiel *Intelligent Essay Assessor*, similairement, crée une représentation du texte en entier, puis calcule la valeur des variables pour le contenu, le style et l'orthographe avant d'attribuer une note basée sur une équation de régression multiple. Les lecteurs entraînés mais peu expérimentés, quant à eux, vont plutôt interpréter une partie du texte, puis l'évaluer, pour ensuite interpréter une autre partie et l'évaluer, divisant ainsi la tâche de correction en de multiples sous-tâches. Ils ne portent donc pas leur jugement sur l'ensemble du texte, mais sur des parties séparées, ce qui peut entraîner un jugement global différent. C'est ce qui expliquerait, selon Nichols, le fait que les lecteurs entraînés ont un taux d'accord entre eux et les correcteurs experts beaucoup moins élevé que le taux d'accord entre le correcticiel et les correcteurs experts.

Comme nous avons pu le constater, le point commun des correcticiels anglophones est leur capacité de corriger à la fois la forme et le sens des textes. Les principales

critiques concernant les correcticiels viennent d'ailleurs du fait qu'ils attribuent une note au sens du texte. Plusieurs craignent que cela envoie un message négatif aux étudiants qui ne seront pas lus par des humains, mais par des machines (Wang et Brown, 2007). D'autres critiques concernent la superficialité des critères de correction sur lesquels se basent les correcticiels ainsi que leur insensibilité au contenu des réponses et à la créativité des élèves (Yang, Buckendahl, Juskiewicz et Bhola, 2002). Les correcticiels anglophones sont donc encore au cœur d'un débat, principalement causé par l'éthique entourant la question de la correction du sens d'un texte par un ordinateur.

Il en va toutefois autrement pour les correcticiels francophones. Les outils de correction comme *Correcteur 101* et *Antidote* ne sont pas en mesure de prendre en compte le sens du texte, mais ils peuvent corriger la forme. Ces types de correcticiels correspondent davantage à ceux dont nous aurions besoin, au Québec, pour corriger une partie des textes d'élèves. En effet, contrairement aux épreuves d'écriture administrées aux États-Unis, les examens du MELS sont corrigés, sur le plan de la forme, de façon détaillée, avec plusieurs critères à évaluer. Les correcteurs ne doivent pas attribuer uniquement une note de 1 à 6 pour l'ensemble du texte. Nous avons donc besoin de correcticiels qui peuvent offrir une correction détaillée plutôt que d'attribuer une note globale. C'est ce que permettent les correcticiels comme *Correcteur 101* et *Antidote*. Voici, avant d'examiner plus attentivement ces deux correcticiels, les principales fonctions et le processus de correction des correcticiels en général.

2.2.3 Fonctions des correcticiels

Dans son article *Correcteurs orthographiques et grammaticaux : Quel(s) outil(s) pour quel rédacteur?*, Jacquet-Pfau (2001) souligne que le principal rôle de l'outil de correction est de détecter, selon des critères préétablis, les erreurs dans un texte. Lorsque le correcticiel détecte une erreur, il peut proposer différentes

corrections. C'est ensuite au lecteur ou au correcteur humain de choisir de remplacer ou non le segment détecté par une des propositions du correcticiel. Les outils de correction automatisée ne corrigeront donc pas, au sens propre, le texte, mais ils détecteront plutôt les erreurs en proposant par la suite certaines corrections à apporter au texte.

2.2.4 Processus de correction

Le processus de correction des correcticiels se décompose en quatre étapes (Jacquet-Pfau, 2001) :

- 1- la vérification lexicale,
- 2- la proposition et le choix du mot à retenir,
- 3- la vérification grammaticale,
- 4- la proposition et le choix de solution.

Le logiciel commence par effectuer une vérification lexicale, pour ensuite faire des propositions de mots à retenir ou à modifier. Selon Désilets (1999), la vérification lexicale, pour un correcticiel, consiste à « s'assurer que chacun des mots du texte ou de la phrase correspond à une forme existante du français » (Désilets, 1999, p. 30). Lorsque le correcticiel détecte une forme lexicale inconnue, il peut proposer une liste de graphies possibles pour modifier la forme fautive. Cette liste est construite par ressemblance phonétique ou alphabétique, selon le type de correcticiel. Par exemple, si le correcticiel détecte l'erreur **farmacie*, il pourra l'associer phonétiquement avec sa forme adéquate et proposer la graphie *pharmacie* (Morin, 1995). Par la suite, le correcticiel effectue une vérification de l'orthographe grammaticale et propose finalement des solutions lorsque des erreurs ont été détectées. Pour l'analyse syntaxique, le correcticiel doit vérifier si les catégories syntaxiques des mots sont compatibles et si l'ordre dans lequel ils sont présents dans la phrase correspond à un des modèles syntaxiques qu'il connaît (Désilets, 1999). Les correcticiels sont en effet

programmés à partir d'une série de modèles syntaxiques. Lorsqu'ils analysent une phrase, ils associent chaque mot à la catégorie grammaticale à laquelle il appartient et vérifient si l'ordre des mots correspond à un modèle syntaxique qu'ils connaissent.

2.2.5 Concept d'erreur dans le cadre d'un correcticiel

Selon Jacquet-Pfau (2001), pour un correcticiel, une erreur consiste en « tout ce qui n'est pas reconnu par le système ». Lorsqu'une chaîne de caractères ne correspond à aucune forme pour laquelle le système a été programmé, une erreur est détectée. Il y a donc détection d'erreur lorsque les formes graphiques ne correspondent pas à une forme mémorisée dans le lexique du système, lorsque les formes graphiques ne correspondent pas aux formes générées automatiquement, soit les formes fléchies, ou lorsque la structure de la phrase ne correspond à aucune structure enregistrée dans la mémoire du correcticiel.

2.2.6 Classification des erreurs

Dans son article, Jacquet-Pfau (2001) établit une typologie des erreurs détectées par les correcticiels. Elle classe tout d'abord dans la catégorie *erreurs d'orthographe* les erreurs relevant soit du niveau lexical, soit du niveau flexionnel, celui-ci se rapportant à la variation d'un mot en contexte syntaxique. Le niveau lexical concerne « toutes les erreurs portant sur l'écriture du mot hors contexte » (Jacquet-Pfau, 2001). Il peut donc s'agir d'erreur au niveau d'une lettre, d'une syllabe, d'un graphème, d'un morphème, d'un mot, d'un doublon (deux mots identiques consécutifs), des marques de genre et de nombre qui ne sont pas des morphèmes admis, comme **chevals* ou des marques de flexion verbale qui ne sont pas des morphèmes admis, comme **disez*. Les règles d'accord et de conjugaison, quant à elles, font partie de la catégorie d'analyse syntaxique. Il s'agit de la classification choisie par Jacquet-Pfau, mais elle peut ne pas correspondre avec la

classification des erreurs dans les différents correcticiels. Ce sont des humains qui programment les correcticiels et ce sont eux qui déterminent la classification des erreurs. Elle peut donc varier d'un correcticiel à l'autre, tout comme elle peut varier d'une grammaire à l'autre et d'un correcteur humain à l'autre.

2.2.7 Paramétrages de la correction

Les systèmes de correction automatisée peuvent offrir différentes modalités de correction. Il est donc possible de demander au correcticiel de concentrer son analyse sur certains aspects ou sur certaines parties du texte. Un logiciel comme *Antidote*, par exemple, peut également proposer différents niveaux d'information. *Antidote* comporte le niveau « alerte » qui ne fait que signaler l'erreur à l'aide d'un bref message au-dessus de la détection effectuée, le niveau « explication contextuelle », qui fournit un message détaillé dans une nouvelle fenêtre, et le niveau « grammaire », qui fait apparaître la règle appliquée. Il est donc possible de paramétrer la correction afin qu'elle corresponde aux besoins des utilisateurs. Les correcteurs experts qui veulent rapidement connaître le nombre d'erreurs dans un texte pourront utiliser un paramétrage différent de l'enseignant qui veut aider ses élèves en leur enseignant comment réviser leurs textes et corriger leurs erreurs à l'aide d'outils de correction.

2.2.8 Avantages et limites

Les correcticiels présentent l'avantage de ne pas subir, contrairement aux correcteurs humains, d'influence externe : ils ne peuvent être soumis aux facteurs psychologiques et physiologiques comme le stress ou la fatigue (Cordier-Gauthier et Dion, 2003). Toutefois, les correcticiels présentent certaines limites. Contrairement aux correcteurs humains, ils ne sont pas des « lecteurs » (Cordier-Gauthier et Dion, 2003). En effet, les correcticiels ne peuvent véritablement tenir compte du sens des textes qu'ils corrigent. Il peut donc arriver qu'ils effectuent des diagnostics inexacts

en analysant les différents éléments de la phrase sans tenir compte du contexte phrastique et discursif. Les solutions qu'ils proposent ne sont alors pas toujours adéquates. Les correcticiels peuvent également être sujets à des erreurs de système que Jacquet-Pfau (2001) nomme « silence » ou « bruit ». Le silence consiste à omettre de signaler une erreur. Au niveau de la correction orthographique, ce type d'erreur de détection peut concerner les erreurs d'accord, les homonymes ou tout mot figurant dans un dictionnaire. Un correcticiel pourrait, par exemple, ne pas signaler l'erreur dans la phrase **Je mange des amendes* s'il ne tient pas compte du contexte dans lequel le mot *amende* doit se retrouver. Il pourrait également omettre de signaler l'erreur dans la phrase **Les enfants de la fille de mon cousin rentre de Paris* si les écrans au singulier *la fille* ou *mon cousin* l'empêchent d'effectuer correctement l'accord du verbe avec la tête du GN sujet, *les enfants*, qui est au pluriel. Jacquet-Pfau (2001) suggère, pour pallier cette limite, d'intégrer aux correcticiels un lexique plus complet et davantage de règles grammaticales. La seconde faille du système, le « bruit », consiste en une détection d'erreur alors qu'il n'y en a pas. Un correcticiel pourrait indiquer une erreur dans la phrase *Les enfants de mon voisin sont chez moi* parce qu'il aurait effectué l'accord avec *voisin* plutôt qu'avec *les enfants*. Cette faille du système concerne surtout la vérification syntaxique. En effet, il semble qu'il s'avère difficile actuellement pour des correcticiels comme *Correcteur 101* et *Antidote* d'analyser des phrases longues et complexes et ils auront tendance à détecter toutes les erreurs possibles (Jacquet-Pfau, 2001). Outre les phrases longues et complexes, les correcticiels ont également du mal à analyser les phrases qui sont mal construites, c'est-à-dire celles présentant des erreurs importantes de syntaxe, ce qui risque de poser problème lors de la correction de textes d'élèves du premier cycle du secondaire.

2.3 Comparaison entre deux correcticiels francophones

Le *Correcteur 101* et *Antidote* sont, à l'heure actuelle, les deux correcticiels francophones disponibles au Québec. Il existe également des correcteurs intégrés à

certain logiciels de traitement de texte, comme *Microsoft Word*, mais nous avons préféré porter notre attention sur des logiciels conçus spécifiquement pour la correction des textes plutôt que sur des logiciels de traitement de texte offrant un outil complémentaire de correction. C'est pourquoi nous traiterons uniquement, dans cette section, du *Correcteur 101* et d'*Antidote*. La comparaison de ces deux correcticiels nous aidera à choisir lequel nous retiendrons pour cette recherche.

2.3.1 *Correcteur 101*

Certains chercheurs se sont penchés sur le fonctionnement et l'efficacité du *Correcteur 101*. Le terme efficacité⁸ renvoie ici au nombre d'erreurs correctement détectées par le correcticiel dans un texte, par rapport au nombre d'erreurs total. Au niveau de ses caractéristiques techniques, Morin a noté que la version de 1992 du correcticiel peut repérer 1 500 difficultés de la langue française. Il peut également reconnaître les paronymes et les homophones (Morin, 1995). Sa grammaire interne comporte 2 500 règles syntaxiques et 300 règles d'accord. Son dictionnaire intégré renferme 60 000 mots. Ce dictionnaire est inspiré partiellement du lexique-grammaire de Maurice Gross de l'Université Paris III. La principale lacune du *Correcteur 101*, selon Morin, est qu'il analyse les phrases une à une. Il ne tient donc compte ni de la phrase précédente, ni de la suivante. Il ne peut en conséquence pas vérifier les éléments liés à la cohérence du texte, soit la concordance des temps verbaux, les reprises, l'utilisation correcte des connecteurs, les répétitions abusives, etc.

Désilets, en 1999, s'est intéressé à l'efficacité de la version préliminaire du *Correcteur 101 didactique*. Il a soumis le logiciel à 24 textes produits par 9 élèves, soit 3 en sixième année du primaire, 3 en première secondaire et 3 en deuxième secondaire. Dans chacun de ces groupes, un élève était considéré fort (A), un autre moyen (B) et un faible (C). Cette classification avait été établie par l'enseignant. Les 24 textes contenaient, ensemble, 527 erreurs portant sur l'orthographe. 75 % de ces

⁸ Cette définition est celle qui sera retenue pour cette recherche.

erreurs, soit 397 erreurs, ont pu être repérées et classifiées correctement par *Correcteur 101*. 33 erreurs additionnelles, soit 6%, ont été signalées, mais mal classifiées. La performance totale du système, pour le signalement des erreurs, était donc de 81 % dans cette étude. Le tableau présenté ci-dessous (*Tableau II*) précise la performance du correcticiel selon différentes catégories, soit *orthographe lexicale* (LX), *accord en genre et en nombre dans le groupe nominal* (GN), *accord du verbe conjugué* (VC), *accord du participe ou de l'attribut* (PA) et *confusion homophonique* (HM). Les erreurs les plus faciles à repérer, selon ce tableau, sont celles d'orthographe lexicale, avec un taux de repérage total de 96 %. Celles qui sont les plus difficiles à repérer pour le correcticiel sont les erreurs concernant la confusion homophonique, avec un taux de repérage total de 55 %.

Tableau II
Performances du *Correcteur 101**

	Total	LX	GN	VC	PA	HM
Nombre d'erreurs réelles	527	200	100	67	31	129
% du total	100 %	38 %	19 %	13 %	6 %	24 %
Nombre d'erreurs repérées	430					
Nombre d'erreurs repérées bien classées	397	189	76	55	24	53
Nombre d'erreurs repérées mal classées	33	4	5	4	2	18
Nombre d'erreurs non repérées	97	7	19	8	5	58
% d'erreurs repérées bien classées	75 %	94 %	76 %	82 %	77 %	41 %
% d'erreurs repérées mal classées	6 %	2 %	5 %	6 %	6 %	14 %
% d'erreurs non repérées	18 %	4 %	19 %	12 %	16 %	45 %

*Performances du prototype mesurées sur 24 textes d'élèves (pourcentages arrondis)

Source : Désilets, M. (1999). *L'utilisation pédagogique d'un logiciel correcteur au primaire et au secondaire*. Thèse de Doctorat, Université de Montréal, Montréal, p. 65.

L'étude de Désilets portait également sur le nombre de fausses alertes, ce que nous avons précédemment nommé « bruit », commises par le correcticiel. Le tableau présenté à la page suivante (*Tableau III*) montre que *Correcteur 101* a émis un total

de 19 % de fausses alertes. Ces dernières concernaient surtout l'accord des verbes conjugués (31 % des cas) et la confusion homophonique (30 % des cas). Cette version du système émet très rarement de fausses alertes pour des erreurs d'orthographe lexicale (2 % des fausses alertes).

Tableau III
Fausses alertes commises par *Correcteur 101**

	Total	LX	GN	VC	PA	HM
Nombre de signalements	529	196	111	86	35	101
Nombre de fausses alertes	99	3	30	27	9	30
% de fausses alertes	19 %	2 %	27 %	31 %	26 %	30 %

*Nombre de fausses alertes commises par le prototype sur le corpus de 24 textes (pourcentages arrondis)

Source : Désilets, M. (1999). *L'utilisation pédagogique d'un logiciel correcteur au primaire et au secondaire*. Thèse de Doctorat, Université de Montréal, Montréal, p. 66.

Les études sur le *Correcteur 101* effectuées par Morin et par Désilets, bien qu'elles fournissent des renseignements détaillés sur la performance du correcticiel, concernent des versions plus anciennes du logiciel. On peut supposer que le correcticiel n'a pu qu'améliorer son efficacité, mais nous n'avons pas trouvé d'études récentes qui confirmeraient cette hypothèse.

2.3.2 *Antidote*

La principale étude sur *Antidote* est celle de Burston, en 1999. Il indique que la version 98 de ce logiciel comprend un vérificateur grammatical, un dictionnaire, un conjugateur de verbe et une grammaire de référence. Le vérificateur grammatical peut repérer 800 erreurs et est sensible aux homophones, aux paronymes, aux anglicismes et aux régionalismes. Le correcticiel est programmé à partir de la grammaire *Le Bon Usage* de Grevisse. Le dictionnaire intégré contient 100 000 entrées incluant 35 000 définitions tirées du *Dictionnaire de l'Académie française* et de près de 40 autres

sources. Le conjugateur de verbe est doté d'un inventaire de 8 000 verbes sous 400 000 formes conjuguées. Tout comme le *Correcteur 101*, *Antidote* ne peut prendre en compte le sens des phrases qu'il analyse.

Dans son étude, Burston a voulu comparer l'efficacité du *Correcteur 101* (version 3.5 Pro) et d'*Antidote* (version 98). Il a donc soumis les deux correcticiels à différents tests. Le premier test consistait à analyser un extrait de texte littéraire, soit les sept premières phrases du texte *Le Capitaine Fracasse* de Théophile Gautier⁹. Ce texte ne contenait aucune erreur. Les résultats sont regroupés dans le tableau ci-dessous (*Tableau IV*). *Antidote* a émis plus d'analyses partielles, c'est-à-dire des analyses incomplètes des phrases à cause de la complexité de celles-ci, mais moins de commentaires douteux et de fausses corrections. Son temps de correction par phrase était également beaucoup plus rapide.

Tableau IV

Analyse des textes littéraires par les correcticiels

	Analyses partielles	Commentaires douteux	Fausse corrections	Temps maximal par phrase
<i>Correcteur 101</i>	1	8	13	6 minutes
<i>Antidote</i>	3	1	11	4 secondes

Source: Burston, J. (1999). *Antidote 98*. CALICO [version électronique]. Récupéré le 31 mars 2008 de [https://www.calico.org/p-72-Antidote%2098%20\(11999\).html](https://www.calico.org/p-72-Antidote%2098%20(11999).html)

Le deuxième test consistait à faire l'analyse d'un extrait d'article de journal intitulé *Justice pour la justice* (Robert Badinter, *Le Monde*, 8 février 1997). Les résultats obtenus lors de ce test étaient un peu différents de ceux obtenus précédemment. On peut voir, dans le tableau présenté à la page suivante (*Tableau V*), qu'*Antidote* a émis moins d'analyses partielles et de commentaires douteux que *Correcteur 101*, mais davantage de fausses corrections.

⁹ On peut se questionner sur le choix de Burston quant à ce texte de Gautier datant de 1863, qui est écrit dans une forme de français plus ancienne et pour laquelle les correcticiels modernes n'ont peut-être pas été programmés...

Tableau V
Analyse d'un article de journal par les correcticiens

	Analyses partielles	Commentaires douteux	Fausse corrections
<i>Correcteur 101</i>	12	6	6
<i>Antidote</i>	7	3	8

Source: Burston, J. (1999). *Antidote* 98. CALICO [version électronique]. Récupéré le 31 mars 2008 de [https://www.calico.org/p-72-Antidote%2098%20\(11999\).html](https://www.calico.org/p-72-Antidote%2098%20(11999).html)

Enfin, comme troisième test, Burston a soumis aux deux correcticiens 40 textes d'étudiants de deuxième année universitaire contenant 1262 erreurs morpho-syntaxiques. *Antidote* a détecté un total de 82 % d'erreurs, comparativement à 91 % pour le *Correcteur 101* (Tableau VI). Le *Correcteur 101* a toutefois effectué 49 fausses corrections, comparativement à 28 pour *Antidote*. Le *Correcteur 101* a laissé passer 7 % d'erreurs, tandis qu'*Antidote* en a laissé passer 11 %.

Tableau VI
Analyse des compositions d'élèves par les correcticiens

	Nombre d'erreurs	Erreurs corrigées	Erreurs signalées	Total d'erreurs détectées	Erreurs mal identifiées	Erreurs non détectées	Nombre de fausses erreurs
<i>Correcteur 101</i>	1262	74 %	17 %	91 %	7 %	2 %	49
<i>Antidote</i>	1262	77 %	5 %	82 %	11 %	7 %	28

Source: Burston, J. (1999). *Antidote* 98. CALICO [version électronique]. Récupéré le 31 mars 2008 de [https://www.calico.org/p-72-Antidote%2098%20\(11999\).html](https://www.calico.org/p-72-Antidote%2098%20(11999).html)

Cette étude de Burston ne nous permet pas de porter un jugement précis sur l'efficacité relative des deux correcticiens pour la correction de textes d'élèves du secondaire. Le choix du correcticien qui sera retenu pour notre recherche devra donc se faire à partir d'autres critères (voir la section 3.3).

2.4 Bilan et question de recherche

Tel que mentionné précédemment, l'efficacité de divers correcticiels utilisés aux États-Unis a été démontrée. Certains correcticiels francophones, tels *Correcteur 101* et *Antidote*, bien qu'ils fonctionnent différemment des correcticiels anglophones, ont également montré un taux d'efficacité élevé pour détecter les erreurs dans des textes. Toutefois, les études sur les correcticiels francophones datent de plusieurs années et des versions plus récentes de ces correcticiels sont aujourd'hui disponibles. Nous sommes donc intéressée à connaître les différences entre la correction humaine et la correction effectuée par une version récente d'un correcticiel francophone lors de la correction de textes d'élèves du secondaire. Ainsi, nous voulons observer les différences quant au nombre et à la classification des erreurs repérées. Nous voulons également connaître l'efficacité des versions récentes des correcticiels lorsqu'ils sont utilisés pour corriger des textes d'élèves du secondaire. Voici donc, en résumé, l'objectif, la question générale de cette recherche et les questions spécifiques :

Objectif : Comparer la correction effectuée par des correcteurs humains à celle effectuée par un correcticiel lors d'une épreuve de français.

Question générale : Quelles sont les différences entre la correction humaine et la correction effectuée par une version récente d'un correcticiel francophone lors de la correction de la qualité de la langue dans des textes d'élèves du secondaire?

Questions spécifiques :

- Quelles sont les différences quant au nombre d'erreurs détectées?
- Quelles sont les différences quant à la classification des erreurs?
- Quelle est l'efficacité des versions récentes des correcticiels lorsqu'ils sont utilisés pour corriger des textes d'élèves du secondaire?

Le prochain chapitre présentera la méthodologie que nous avons suivie afin de répondre à ces questions de recherche.

Troisième chapitre

Méthodologie de la recherche

3.1 Type de recherche

Selon Boudreault (2004), la recherche scientifique, dite expérimentale, « repose sur la démonstration de liens significatifs de cause à effet entre différentes variables » (Boudreault, 2004, p. 152). Dans la présente recherche, nous avons voulu observer l'effet que la variable indépendante, soit la modalité de correction (humaine ou informatisée), a sur la variable dépendante, c'est-à-dire les résultats de la correction. Cette recherche est de type quasi-expérimental puisque l'échantillonnage ne s'est pas fait de façon aléatoire (Boudreault, 2004). En effet, nous avons plutôt opté pour un échantillonnage intentionnellement critérié avec des textes choisis, tel qu'il sera décrit dans la prochaine section, en fonction des notes leur ayant été attribuées.

3.2 Élaboration du corpus de données

Les données utilisées pour cette recherche ont été recueillies en deux étapes. Nous avons tout d'abord eu accès à des textes d'élèves de deuxième année du premier cycle du secondaire provenant de trois écoles¹⁰. Ces élèves avaient déjà rédigé un texte descriptif en février 2007 dans le cadre d'une épreuve administrée par le ministère de l'Éducation, du Loisir et du Sport (MELS). Cette épreuve était un prototype d'instrument d'évaluation pouvant servir au bilan de fin de cycle ou d'année. Il visait à soutenir le milieu scolaire pour que ce dernier s'approprie le

¹⁰ Soulignons qu'un certificat d'éthique a été émis pour cette recherche par les autorités responsables de cette émission à l'Université de Montréal.

Programme et qu'il puisse mettre en place des pratiques pédagogiques visant le développement des compétences. Les textes rédigés par les élèves ont été corrigés par des enseignantes recrutées par le MELS et, ensuite, par quatre correcteurs professionnels du MELS. Chaque texte était donc corrigé deux fois : une première fois par une enseignante, puis une seconde fois par un des correcteurs du MELS. Les correcteurs du MELS devaient corriger chaque texte en soulignant les erreurs et en indiquant dans la marge la catégorie à laquelle chacune de ces erreurs appartenait, soit « U » pour *orthographe d'usage*, « G » pour *orthographe grammaticale*, « P » pour *punctuation* et « S » pour *syntaxe*. De tous les textes corrigés, 142 ont été mis à notre disposition. Ces 142 textes avaient déjà été choisis par les responsables de l'épreuve du MELS. Ils avaient retenu 48 textes avancés, 46 moyens et 48 faibles. De ces 142 textes, seulement 45 ont été retenus pour la recherche. Ce nombre semble suffisant puisque Burston, en 1999, avait soumis seulement 40 textes au logiciel *Antidote* afin d'étudier sa façon de corriger. Notre étude étant plus qualitative que celle de Burston, notre échantillon de 45 textes nous fournira suffisamment de matériel pour analyser le fonctionnement du correcticiel choisi pour notre étude. Parmi les 45 textes, nous en avons choisi 15 très bons, c'est-à-dire se retrouvant dans les échelons de correction 4 ou 5 pour les critères tenant compte de la langue; 15 autres textes étaient moyens, c'est-à-dire dans l'échelon 3, et les 15 derniers textes étaient considérés comme faibles, c'est-à-dire dans les échelons 1 ou 2. Les 15 textes retenus dans chacune des catégories ont été choisis en fonction du type d'erreurs qu'ils contiennent. Pour ce faire, nous avons élaboré une grille¹¹ recensant les différents types d'erreurs observables, que ces dernières fassent partie ou non des erreurs devant être corrigées par le MELS. Nous avons rempli cette grille en examinant les erreurs, détectées ou non par les correcteurs, de chaque texte. Nous avons ainsi tenté d'obtenir un ensemble de textes contenant des erreurs de toutes sortes, soit la construction des phrases passives, interrogatives ou négatives, les participes passés, les erreurs dans le choix de l'auxiliaire ou dans le choix du pronom relatif, etc. Les textes présentant une plus grande variété d'erreurs ont été retenus, de même que ceux contenant des erreurs peu fréquentes dans l'ensemble des textes, afin

¹¹ Voir *Annexe 2*

de mettre le correcticiel à l'épreuve. Cette façon de choisir les textes pour cette recherche rejoint celle de Désilets qui, en 1999, avait utilisé 24 textes pour étudier le fonctionnement du *Correcteur 101*. Ces textes provenaient de 9 élèves, dont 3 étaient considérés forts, 3 moyens et 3 faibles. L'échantillonnage est donc intentionnel critérié, puisque nous choisissons nous-mêmes des textes en fonction des notes qui leur ont été attribuées. Cette façon de sélectionner les textes offre l'avantage de mettre le correcticiel à l'épreuve sur une plus grande panoplie d'erreurs, les forts ne faisant peut-être pas les mêmes erreurs que les faibles. Soulignons que le choix des 45 textes retenus aurait tout aussi bien pu être différent, et ce, en conservant les mêmes critères de sélection. L'important, pour cette recherche, était de sélectionner une variété de textes de forces différentes présentant des erreurs variées.

La seconde étape de l'élaboration du corpus de données a été de soumettre les textes d'élèves sélectionnés au correcticiel. Pour ce faire, les textes écrits à la main ont dû être retranscrits à l'ordinateur par une personne engagée à cet effet par le directeur de notre équipe de recherche (le GRIÉMÉtic). Ce processus de cueillette de données avait été utilisé par Burstein, Kukich, Wolff, Lu et Chodorow, en 1996. Afin d'éviter des biais possibles, nous avons dû relire les textes retranscrits à l'ordinateur et nous assurer qu'ils soient identiques, en ce qui concerne les erreurs présentes, à ceux rédigés à la main. Certaines corrections ont dû être apportées, soit en moyenne 1 ou 2 par texte. En effet, lors de la retranscription, certaines erreurs avaient été involontairement corrigées, tandis que des erreurs de frappe s'étaient insérées à quelques endroits.

3.3 Description du fonctionnement du correcticiel

Pour la correction automatisée des textes, le correcticiel qui a été utilisé est *Antidote*. Nous avons choisi ce correcticiel parce qu'il est relativement récent et que peu d'études se sont penchées sur son fonctionnement. C'est un outil de plus en plus populaire qui semble offrir de nouvelles options de correction intéressantes. Nous

avons utilisé la toute dernière version à ce jour, soit *Antidote RX*. Cette version permet, selon le guide d'utilisation, de traiter certains aspects pragmatiques (Qui, Quand, Où, Combien), logiques (charnières, citations), stylistiques (répétitions, verbes ternes¹², tournures passives et négatives) et sémantiques (positif, négatif, fort, faible). Le processus de correction d'*Antidote RX* est également très intéressant. En effet, le correcticiel propose différents types de soulignement des erreurs en fonction de la nature de l'erreur commise :

- Le correcticiel souligne les erreurs les plus importantes en rouge;
- l'orangé est utilisé pour les erreurs possibles;
- le souligné rouge plein indique une erreur majeure que le correcticiel peut corriger seul;
- le souligné rouge pointillé signale une erreur majeure qui doit être corrigée par l'utilisateur du correcticiel;
- le souligné orange gras signale une erreur mineure, comme un espacement ou un guillemet, que le correcticiel peut corriger seul;
- le souligné orangé fin, quant à lui, est une alerte, c'est-à-dire qu'il indique une erreur potentielle;
- un trait vertical orangé indique une discontinuité, c'est-à-dire que le correcticiel ne peut analyser complètement la phrase, généralement à cause de la complexité de cette dernière.

La figure¹³ présentée à la page suivante (*Fig. 2*) illustre les différents soulignements utilisés lors de la correction par *Antidote*.

¹² Avoir, être, faire, etc.

¹³ La reproduction des captures d'écran est autorisée par *Druide informatique inc.* Une copie de l'autorisation se trouve en annexe (*Annexe 3*).

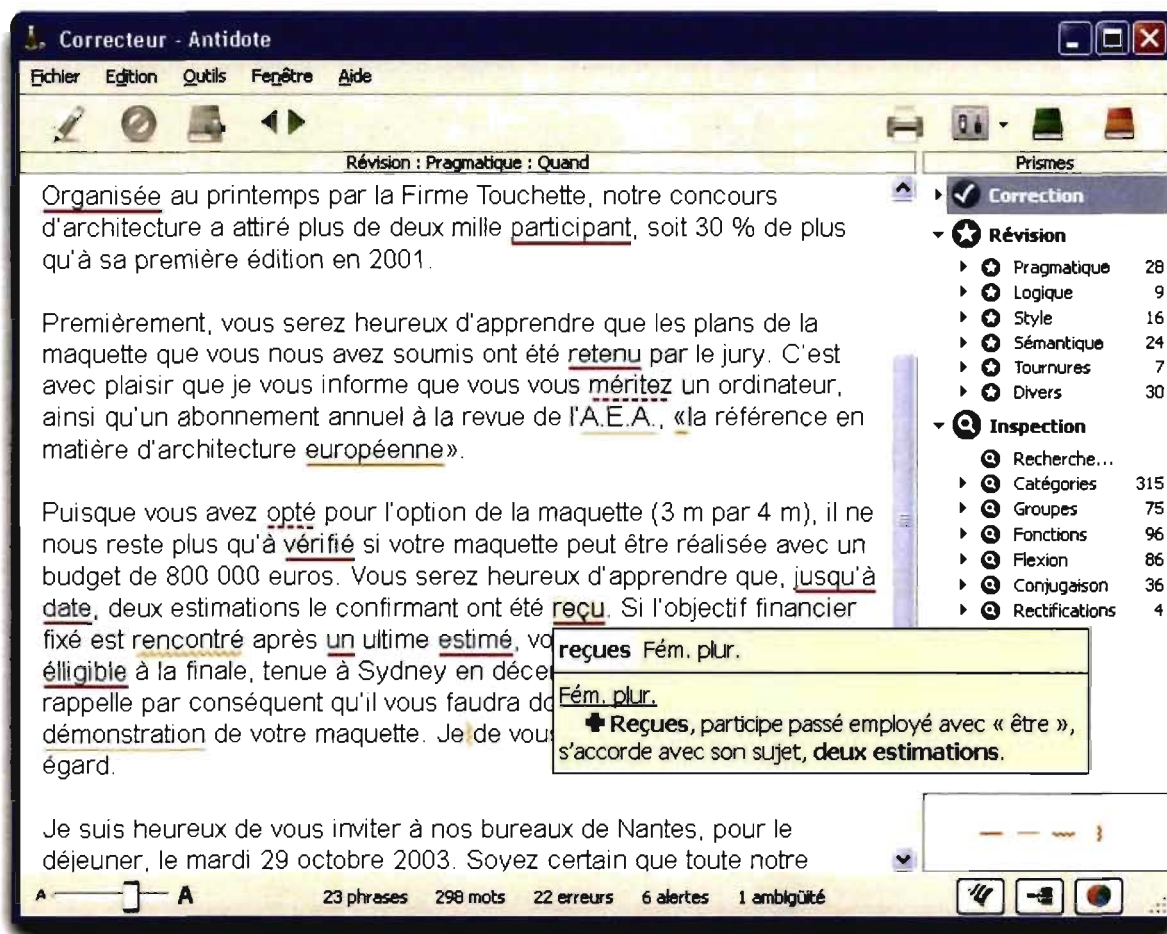


Fig. 2 - Écran de correction d'Antidote

Source : Druide (2008). *Antidote.info* [version électronique]. Récupéré le 2 février 2009 de http://www.druide.com/a_correcteur.html

Outre les erreurs soulignées par le correcticiel, une propriété importante de la dernière version d'*Antidote* est son prisme de correction. Le guide d'utilisation indique qu'une « quarantaine de filtres de correction isolent les corrections du correcteur selon la nature des erreurs : Orthographe, Lexique, Grammaire¹⁴, Syntaxe, Ponctuation, Rédaction et Typographie » (Druide, 2006, p. 33). *Antidote* peut également fournir des statistiques de correction. Un onglet nommé *Erreurs* (Fig. 3, page suivante)

¹⁴ Le terme *grammaire* renvoie, dans ce contexte, à l'orthographe grammaticale.

regroupe les statistiques¹⁵ sur la nature et le nombre d'erreurs, tandis qu'un onglet nommé *Contenu* (Fig. 4, page suivante) fournit des informations sur les éléments lexicaux du texte analysé.

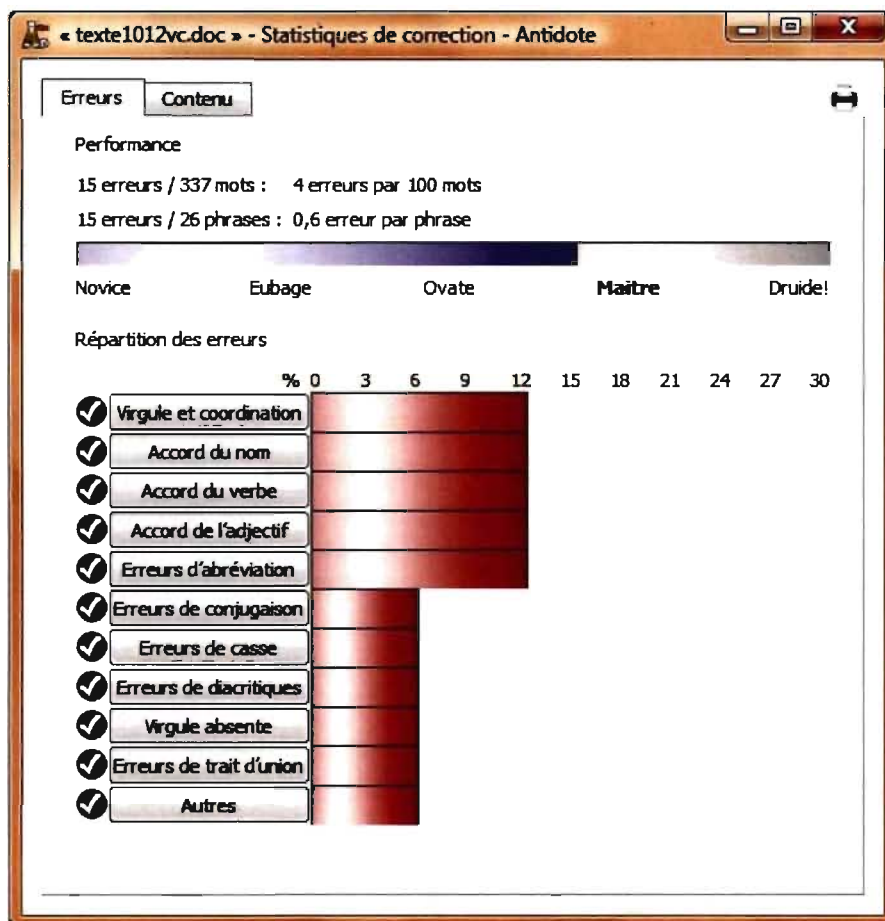


Fig. 3 - *Statistiques de correction – onglet Erreurs*

¹⁵ Précisons que, pour les statistiques d'*Antidote*, le nombre d'erreurs correspond aux erreurs dont il est sûr. Les alertes, les ambiguïtés et les analyses partielles ne sont pas comptabilisées dans ses statistiques. Le calcul du nombre d'erreurs que nous ferons pour notre recherche sera différent, puisque nous ne voulons pas écarter les erreurs potentielles.

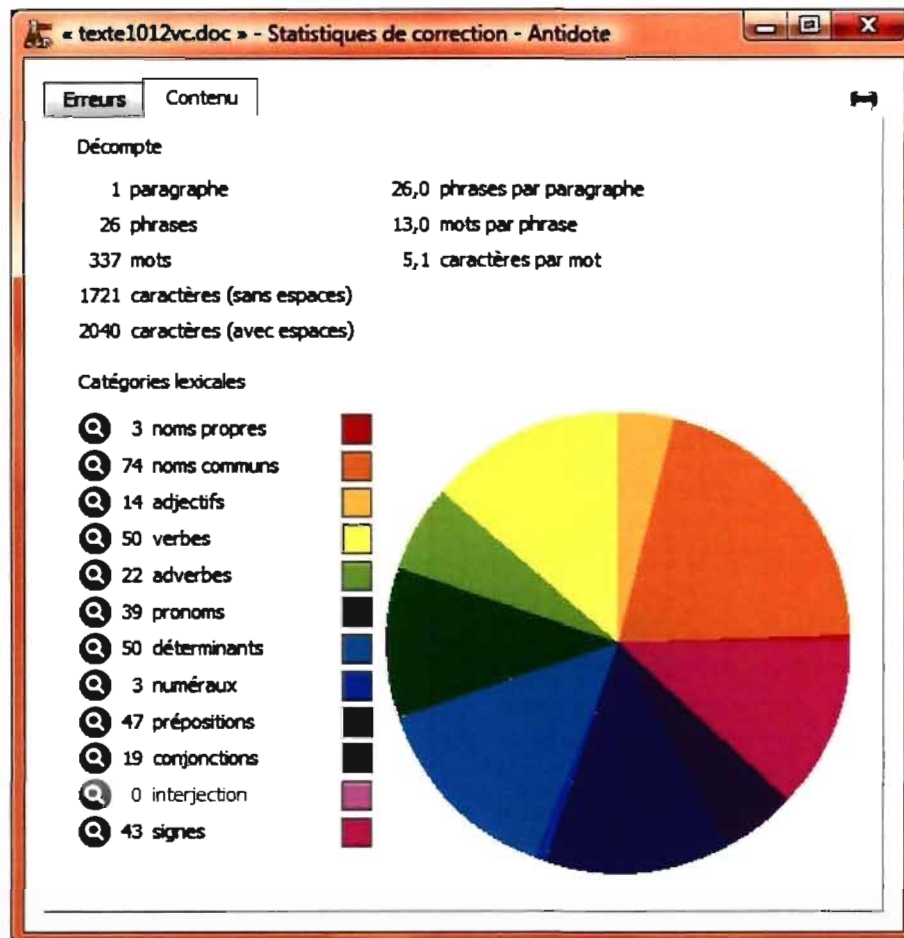


Fig. 4 - *Statistiques de correction – onglet Contenu*

L'outil de statistiques permet également de voir le niveau auquel *Antidote* classe le texte (Fig. 3, page précédente). En effet, 5 niveaux de performance sont déterminés, soit *Novice*, *Eubage*, *Ovate*, *Maître* et *Druide*. *Antidote* classe le texte sur une échelle de niveau à partir du ratio du nombre d'erreurs par 100 mots et du nombre d'erreurs par phrase. Enfin, le correcticiel peut être ajusté à partir de différents réglages sur le niveau de langue, les termes inconvenants ou offensants, les anglicismes, les confusions lexicales, les majuscules, la ponctuation, etc. Cette particularité est très intéressante lorsque les correcteurs ne doivent pas corriger les erreurs relevant d'un certain niveau de difficulté, comme c'est le cas dans cette recherche.

Pour la présente étude, les niveaux de réglage du correcticiel ont été utilisés afin que la correction informatisée soit soumise, pour certaines règles, aux mêmes exigences

de correction que celles des correcteurs humains, soit les exigences fournies par le MELS dans son guide de correction. Le correcticiel ne peut ignorer toutes les erreurs qui ne sont pas, selon les grilles du MELS, relevées dans des textes de 2^e secondaire, mais il peut corriger ou ignorer certaines erreurs relatives au niveau de langue, aux anglicismes et à la ponctuation. Les statistiques de correction ont ensuite été exploitées afin de comparer le niveau de performance attribué par le correcticiel et par les correcteurs humains. Ensuite, le prisme de correction a permis d'observer chacune des erreurs identifiées et classées par *Antidote* selon les différentes catégories. Puis, il a été possible d'observer plus attentivement certaines analyses du texte en se tournant vers les différents éléments soulignés.

3.4 Description des analyses

Une analyse mixte, c'est-à-dire à la fois quantitative et qualitative, a été utilisée pour traiter les données. Le logiciel *QDA-Miner* a été employé afin d'effectuer le codage des textes. Les 45 textes retenus ont été importés dans ce logiciel. Nous avons fait un premier codage en identifiant toutes les erreurs relevées par les correcteurs humains. Les textes ont ensuite été codés une seconde fois avec les erreurs identifiées par le correcticiel. Nous avons vérifié toutes les données en effectuant les deux codages une deuxième fois afin de nous assurer qu'il n'y avait pas eu d'erreur d'inattention. Une fois tous les textes codés, nous avons pu procéder à l'analyse des données.

Au niveau quantitatif, l'analyse consistait à comparer le nombre d'erreurs identifiées par les correcteurs humains et par le correcticiel, sans juger de l'exactitude de l'erreur détectée. Nous avons également comparé le niveau de performance attribué au texte par le correcticiel et par les correcteurs humains. Du point de vue qualitatif, il était intéressant de comparer la classification des erreurs qu'effectuent les correcteurs humains et le correcticiel. Nous avons ainsi examiné la nature des erreurs détectées

par *Antidote*, les erreurs correctement ou incorrectement identifiées par chaque type de correcteur et les différences dans leur façon de classer les erreurs.

3.5 Limites de la recherche

Tel que mentionné précédemment, les textes que nous avons reçus du MELS avaient été corrigés une première fois par des enseignantes, puis par les correcteurs du MELS. Or, il est difficile de savoir si certaines erreurs auraient pu échapper aux correcteurs du MELS si les textes n'avaient pas été corrigés une première fois. Il y a donc un biais lors de la comparaison des résultats accordés par les correcteurs humains et par le correcticiel. Il faut garder en tête que les correcteurs humains ont fait une double correction, ce qui n'est pas courant lors des corrections habituelles des examens de français au Québec dans le cadre des épreuves uniformes du MELS.

Le contexte dans lequel l'épreuve prototypique a été administrée n'est également pas le même que celui des épreuves uniformes du MELS. Lors des examens d'écriture de 5^e secondaire, les correcteurs ont un temps très précis à respecter pour la correction de chaque texte. Les textes que nous avons utilisés pour notre recherche n'ont pas été corrigés dans le même contexte. Les correcteurs ont probablement eu plus de temps pour faire la correction. Est-ce que cela affecte la qualité de la correction en terme du nombre d'erreurs détectées? Peut-être.

De plus, les textes utilisés pour notre recherche ont été corrigés par quatre correcteurs du MELS. Nous n'avons pas cherché à comparer la correction faite par chacun des quatre correcteurs et nous ne pouvons nous assurer qu'ils corrigeaient tous exactement de la même façon. Le but de notre recherche n'était pas de comparer les différents correcteurs humains entre eux, mais nous devons être conscients que la correction des textes aurait peut-être varié avec quatre correcteurs différents.

Une autre limite concerne notre corpus de données. Nous avons reçu 142 textes sélectionnés par le MELS, puis en avons retenu 45 selon différents critères mentionnés précédemment. Nos résultats vont être applicables à notre échantillon, mais nous ne pouvons assurer leur généralisation. Rappelons toutefois que notre objectif de recherche est de comparer la correction effectuée par un correcticiel et par des correcteurs humains, ce que notre échantillon de textes nous permet très bien de faire.

Enfin, les limites de notre recherche touchent également les directives contenues dans le guide de correction du MELS. Les correcteurs devaient suivre certaines consignes en ne corrigeant pas les erreurs relevant d'un certain niveau de difficulté. Ces directives sont propres aux textes d'élèves de 2^o secondaire et les correcteurs du MELS ne suivent pas les mêmes consignes lorsqu'ils corrigent des textes d'un autre niveau. Nous ne pouvons donc pas comparer leur habileté à corriger des erreurs d'un niveau plus élevé avec *Antidote*. *Antidote* avait également été calibré, pour cette recherche, avec certains paramètres concernant le niveau de langue, les anglicismes et la ponctuation. Notre recherche se limite à l'analyse d'*Antidote* selon ces paramètres qui, soit dit en passant, ne correspondent pas exactement aux erreurs devant être mises de côté par les correcteurs humains. En effet, *Antidote* ne peut, par exemple, ne pas corriger les participes passés des verbes pronominaux, qui ne sont pas au programme en deuxième secondaire. Il faut donc garder en tête que le correcticiel peut certes être calibré, mais qu'il y a une différence entre les paramètres des correcteurs humains et ceux d'*Antidote*.

Quatrième chapitre

Résultats des analyses

4.1 Analyse quantitative

Afin d'obtenir une vue d'ensemble des données, nous avons commencé par comparer, de façon générale, le nombre d'erreurs détectées par les deux types de correcteur. Il est important de préciser que le nombre d'erreurs détectées correspond, pour les correcteurs humains, au nombre d'erreurs qu'ils ont identifiées, que ces erreurs aient été comptabilisées ou non dans le résultat de l'élève et que ces erreurs en soient réellement ou non.¹⁶ Pour *Antidote*, le nombre d'erreurs détectées regroupe les erreurs identifiées et classées par le biais du prisme de correction, les ambiguïtés et les analyses partielles. Les alertes faisant déjà partie des erreurs identifiées et classées par le biais du prisme de correction, nous ne les avons pas incluses une seconde fois dans le décompte des erreurs. Nous avons toutefois ajouté les erreurs de style concernant l'absence du verbe. Cet élément se trouve dans l'onglet de révision plutôt que dans celui de correction, mais nous jugions approprié de l'inclure, puisque ce type d'erreur rejoint les erreurs de syntaxe. Enfin, nous n'avons pas comptabilisé les erreurs relevant de l'espacement entre les mots puisque les textes n'ont pas été directement rédigés à l'ordinateur par les élèves et que ces erreurs ne pourraient se retrouver dans des copies manuscrites.

En comparant le nombre d'erreurs détectées par chaque type de correcteur, nous avons pu constater que, de façon générale, les correcteurs humains détectent davantage d'erreurs qu'*Antidote* (voir *Tableau IX*, p. 51). Dans l'ensemble des 45 textes, les correcteurs humains ont détecté 1 704 erreurs, tandis qu'*Antidote* en a

¹⁶ L'analyse de l'identification adéquate ou non des erreurs se fera plus loin dans la recherche (section 4.2). Pour l'instant, nous ne voulons que comparer le nombre d'erreurs identifiées par les correcteurs.

détecté 1 372. Il y a donc un écart de 19,4 % entre les 2 types de correcteur pour le nombre total d'erreurs détectées.

Nous avons également voulu connaître la répartition des erreurs détectées par chaque type de correcteur selon les différentes catégories d'erreurs (Fig. 5, p. 49). Pour les correcteurs humains, nous nous retrouvons avec 6 catégories d'erreurs : *syntaxe* (S), *punctuation* (P), *orthographe d'usage* (U), *orthographe grammaticale* (G), *vocabulaire* (V) et les erreurs identifiées, mais *non classées* (NC). Pour *Antidote*, le classement des erreurs est différent. Nous avons retenu, parmi les catégories d'*Antidote*, les 10 catégories suivantes : *syntaxe* (Antidote-S), *punctuation* (Antidote-P), *orthographe d'usage* (Antidote-O), *orthographe grammaticale* (Antidote-G), *lexique* (Antidote-L), *rédaction* (Antidote-R), *typographie* (T-casse), *style* (Style-Vabsent), *analyse partielle* (AP) et *ambiguïté*. Afin de pouvoir faire une comparaison qui soit la plus juste possible, nous avons regroupé de la façon suivante certaines catégories d'*Antidote* pour qu'elles correspondent davantage à celles du MELS (Tableau VII).

Tableau VII
Correspondance entre les catégories

Catégories	MELS	<i>Antidote</i>
Syntaxe	syntaxe (S)	syntaxe (Antidote-S) style (Style-Vabsent) analyse partielle (AP)
Orthographe d'usage	orthographe d'usage (U)	orthographe d'usage (Antidote-O) rédaction (Antidote-R)
Orthographe grammaticale	orthographe grammaticale (G)	orthographe grammaticale (Antidote-G)
Ponctuation	ponctuation (P)	ponctuation (Antidote-P) typographie (T-Casse)
Vocabulaire	vocabulaire (V)	lexique (Antidote-L)
Non classées	non classées (NC)	
Ambiguïté		ambiguïté

Nous avons tout d'abord regroupé les catégories *syntaxe*, *style* et *analyse partielle*. La catégorie *style* correspond aux phrases, détectées par *Antidote*, où le verbe est absent. La majorité de ces phrases ont été classées dans la catégorie *syntaxe* par le MELS, bien que certaines aient plutôt été classées dans la catégorie *ponctuation*. Nous avons toutefois choisi, pour une comparaison d'ensemble, de joindre la catégorie *style* à celle de *syntaxe*. La phrase suivante est un exemple d'erreur de style, pour *Antidote*, correspondant à une erreur de syntaxe pour le MELS : **Deuxièmement, les conséquences sur les humains*. La catégorie *analyse partielle*, quant à elle, correspond surtout aux phrases qu'*Antidote* n'est pas capable d'analyser, et ceci est principalement causé par une erreur de syntaxe. Une phrase comme **Des maladies iraient même jusqu'à cardiovasculaire et respiratoire*, classée dans la catégorie *analyse partielle* par *Antidote*, devient ainsi, pour notre classement, une erreur de syntaxe. Notre second regroupement concerne la catégorie *ponctuation*. Nous avons décidé de joindre la catégorie *typographie* à celle de *ponctuation*, même si certaines erreurs typographiques sont parfois classées, par le MELS, dans la catégorie *orthographe d'usage*. **L'ouragan katrina* est un exemple typique d'erreur typographique, pour *Antidote*, que nous regroupons avec les erreurs de ponctuation. La catégorie *rédaction*¹⁷ a été groupée avec celle d'*orthographe d'usage* puisque la plupart des erreurs se retrouvant dans cette catégorie concernent des abréviations (exemple : **ect*), erreurs classées dans l'*orthographe d'usage* par le MELS. Encore une fois, ce classement est général et nous verrons en détail, à la section 4.2, les différences de classement entre les deux types de correcteur. Soulignons que la catégorie *ambiguïté* n'a pas été groupée avec d'autres catégories puisque les ambiguïtés identifiées par *Antidote* peuvent concerner l'orthographe d'usage ou l'orthographe grammaticale, la syntaxe ou la ponctuation. Par exemple, dans la phrase **La chaleur est le facteur principale de la fonte des glaciers*, *Antidote* a indiqué une ambiguïté qui correspond en fait à une erreur d'orthographe

¹⁷ La catégorie *rédaction* regroupe « l'ensemble des conventions d'écriture qui s'appliquent dans la rédaction de documents particuliers (lettres administratives, curriculum vitæ, courriels, etc.) ou dans l'écriture d'éléments régis par des normes strictes (adresse postale, vedette de la lettre, date, etc.) » (Druide, 2007)

grammaticale, tandis que dans la phrase **Le niveau d'eau va baisser donc, la production énergétique sera difficile à faire*, le correcticiel a également indiqué une ambiguïté, mais qui correspond cette fois à une erreur de ponctuation. Le tableau présenté à la page 46 (*Tableau VII*) résume les nouvelles catégories créées pour notre comparaison.

Le tableau ci-dessous (*Tableau VIII*), ainsi que le graphique qui l'accompagne (*Fig. 5*, page suivante) illustrent la comparaison du nombre d'erreurs détectées par *Antidote* et par les correcteurs humains pour chacune des catégories.

Tableau VIII

Nombre d'erreurs détectées selon les catégories

Catégories		Type de correcteur	Nombre d'erreurs détectées			Total
Syntaxe	Syntaxe (S)	MELS*	312			312
	Syntaxe (Antidote-S), Style-Vabsent, Analyse partielle (AP)	<i>Antidote</i>	57	86	169	312
Ponctuation	Ponctuation (P)	MELS	331			331
	Ponctuation (Antidote-P), Typographie (T-Casse)	<i>Antidote</i>	108	51		159
Orthographe d'usage	Orthographe d'usage (U)	MELS	322			322
	Orthographe d'usage (Antidote-O), Rédaction (Antidote-R)	<i>Antidote</i>	424	11		435
Orthographe grammaticale	Orthographe grammaticale (G)	MELS	625			625
	Orthographe grammaticale (Antidote-G)	<i>Antidote</i>	365			365
Vocabulaire	Vocabulaire (V)	MELS	71			71
	Lexique (Antidote-L)	<i>Antidote</i>	35			35
Non classées	Non classées (NC)	MELS	48			48
Ambiguïté	Ambiguïté	<i>Antidote</i>	79			79

* Le nombre d'erreurs détectées par le MELS comprend uniquement les erreurs que les correcteurs du MELS ont repérées lorsqu'ils ont fait la correction des textes après les enseignants. Les erreurs que les enseignants ont détectées mais que le MELS n'a pas retenues ne sont pas comptabilisées.

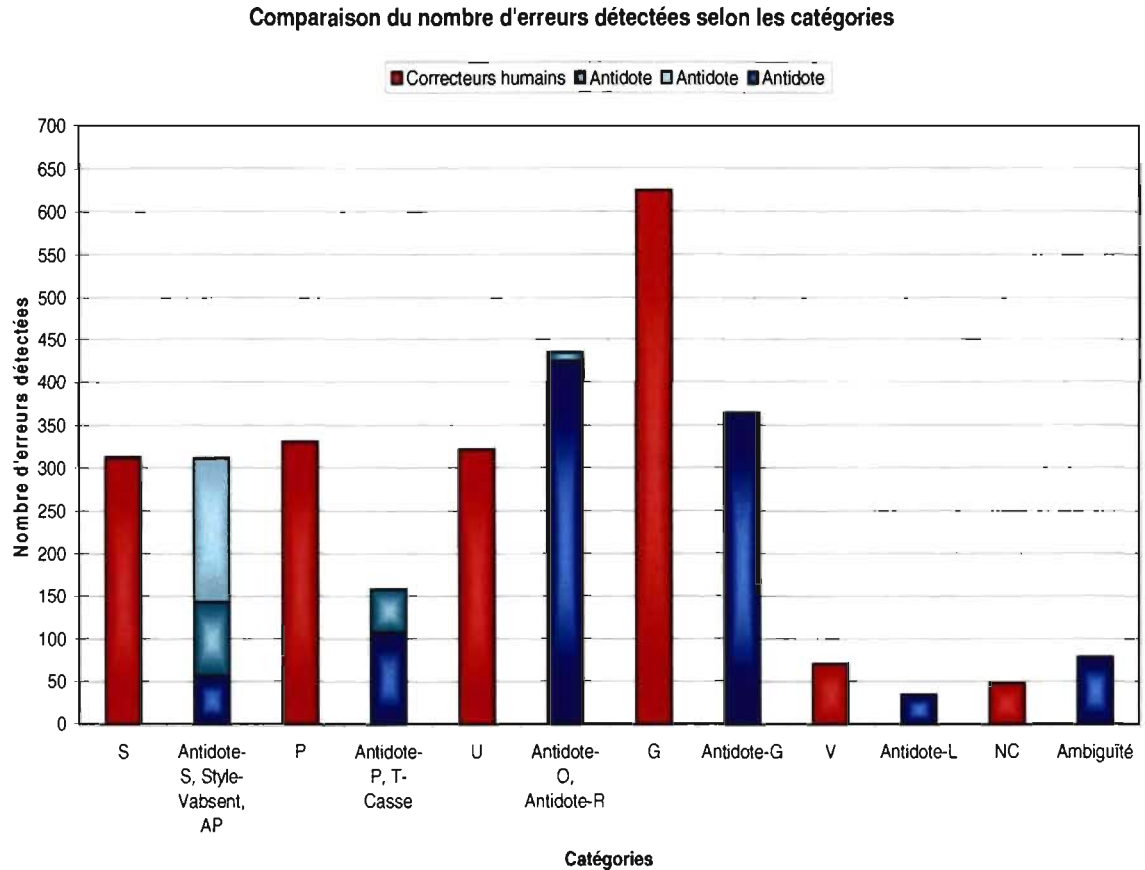


Fig. 5 - *Nombre d'erreurs détectées selon les catégories*

On peut voir, dans la figure 5, qu'*Antidote* a détecté autant d'erreurs de syntaxe (Antidote-S, Style-Vabsent, AP) que les correcteurs humains (S). Il faut toutefois préciser que très peu d'erreurs ont été classées dans la catégorie *syntaxe* par *Antidote*. La plupart des erreurs de syntaxe se sont retrouvées dans les catégories *style-verbe absent* et *analyse partielle*. Nous pouvons également observer que la seule catégorie pour laquelle *Antidote* a repéré plus d'erreurs que les correcteurs humains est l'*orthographe d'usage* (U vs Antidote-O, Antidote-R). Pour les catégories *punctuation* (P vs Antidote-P, T-Casse), *orthographe grammaticale* (G vs Antidote-G) et *vocabulaire* (V vs Antidote-L), *Antidote* a identifié moins d'erreurs que les correcteurs humains.

Maintenant que nous avons une vue d'ensemble du nombre d'erreurs détectées selon le type de correcteur, nous avons voulu vérifier si cette répartition était la même peu importe le niveau du texte. Nous avons donc examiné le nombre d'erreurs selon le type de correcteur en fonction de la variable « niveau du texte » (faible, moyen ou fort) attribuée par les correcteurs du MELS. Dans un premier temps, pour les textes faibles, les correcteurs humains ont repéré 896 erreurs, tandis qu'*Antidote* en a repéré 655 (*Tableau IX*, page suivante). On se retrouve maintenant avec un écart de 26,9 % entre les 2 types de correcteur, ce qui est supérieur à l'écart général pour tous les niveaux de texte confondus, qui était de 19,4 %. Pour les textes moyens, le même tableau nous montre que les correcteurs humains ont détecté 504 erreurs, tandis qu'*Antidote* en a détecté 404. Nous nous retrouvons donc avec un écart de 19,8 % entre les 2 types de correcteur, ce qui est très similaire à l'écart général pour tous les types de textes qui était de 19,4 %. Enfin, pour les textes forts, les correcteurs humains ont identifié 304 erreurs, tandis qu'*Antidote* en a identifié 313 (*Tableau IX*, page suivante), soit 2,9 % de plus que les correcteurs humains.

Tableau IX
Nombre d'erreurs détectées selon le niveau du texte

Type de correcteur	Nombre d'erreurs détectées			
	Textes faibles	Textes moyens	Textes forts	Tous les textes
Correcteurs humains	896	504	304	1 704
<i>Antidote</i>	655	404	313	1 372
% d'écart	26,9 %	19,8 %	2,9 %	19,4 %

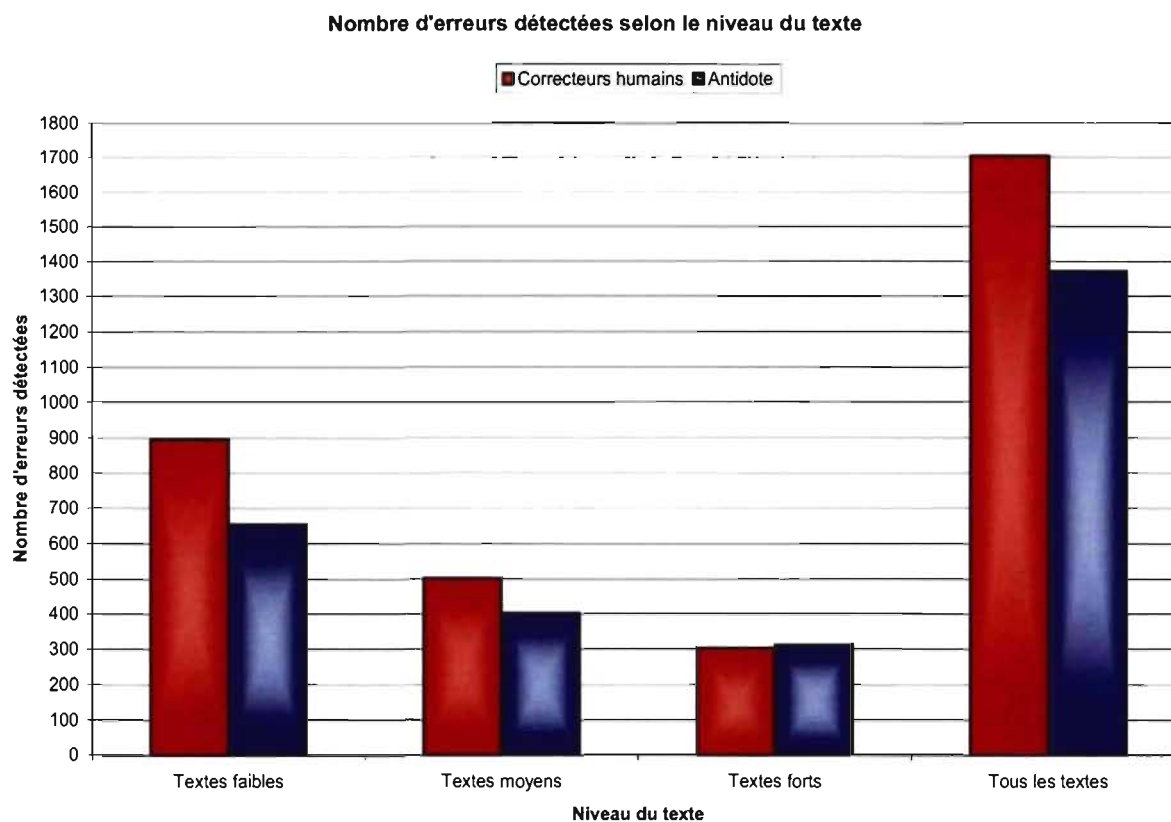


Fig. 6 - Nombre d'erreurs détectées selon le niveau du texte

La répartition des erreurs selon les différentes catégories (Fig. 7-8-9, pages 53-54-55), quant à elle, nous indique tout d'abord qu'*Antidote* identifie plus d'erreurs de syntaxe – style – analyse partielle que les correcteurs humains dans les textes faibles et forts. Plusieurs explications sont possibles : soit *Antidote* identifie des erreurs où il n'y en a pas, soit son analyse est plus poussée que celle des correcteurs humains, qui doivent indiquer uniquement les erreurs du niveau des élèves de deuxième secondaire, ou soit le classement des erreurs est différent entre les deux types de correcteurs. Nous tenterons de trouver l'explication qui s'applique dans la section suivante (4.2). Ensuite, pour les erreurs de ponctuation, nous pouvons observer que, peu importe le niveau du texte, *Antidote* détecte toujours moins d'erreurs que les correcteurs humains. L'écart entre les deux types de correcteur diminue toutefois lorsque le niveau du texte augmente. Pour l'orthographe d'usage et l'orthographe grammaticale, la situation est la même, peu importe le niveau du texte : *Antidote* détecte plus d'erreurs d'orthographe d'usage que les correcteurs humains, mais moins d'erreurs d'orthographe grammaticale. Nous tenterons de vérifier, dans la suite de notre analyse, si cette différence résulte d'un classement différent des erreurs chez les correcteurs. Enfin, en ce qui concerne les erreurs de vocabulaire, *Antidote* en détecte autant que les correcteurs humains dans les textes forts, mais il en détecte moins dans les textes faibles et moyens. Soulignons que le nombre d'erreurs détectées ne signifie pas que les mêmes erreurs ont été identifiées; il peut très bien s'agir d'erreurs différentes. En effet, une erreur de vocabulaire pour les correcteurs humains n'en est peut-être pas une pour le correcticiel et vice-versa.

Tableau X

Nombre d'erreurs détectées selon les catégories – textes faibles

Catégories		Type de correcteur	Nombre d'erreurs détectées			Total
Syntaxe	Syntaxe (S)	MELS	144			144
	Syntaxe (Antidote-S), Style-Vabsent, Analyse partielle (AP)	Antidote	22	31	96	149
Ponctuation	Ponctuation (P)	MELS	178			178
	Ponctuation (Antidote-P), Typographie (T-Casse)	Antidote	43	21		64
Orthographe d'usage	Orthographe d'usage (U)	MELS	197			197
	Orthographe d'usage (Antidote-O), Rédaction (Antidote-R)	Antidote	234	4		237
Orthographe grammaticale	Orthographe grammaticale (G)	MELS	319			319
	Orthographe grammaticale (Antidote-G)	Antidote	163			163
Vocabulaire	Vocabulaire (V)	MELS	31			31
	Lexique (Antidote-L)	Antidote	12			12
Non classées	Non classées (NC)	MELS	29			29
Ambiguïté	Ambiguïté	Antidote	31			31

Comparaison du nombre d'erreurs détectées selon les catégories - textes faibles

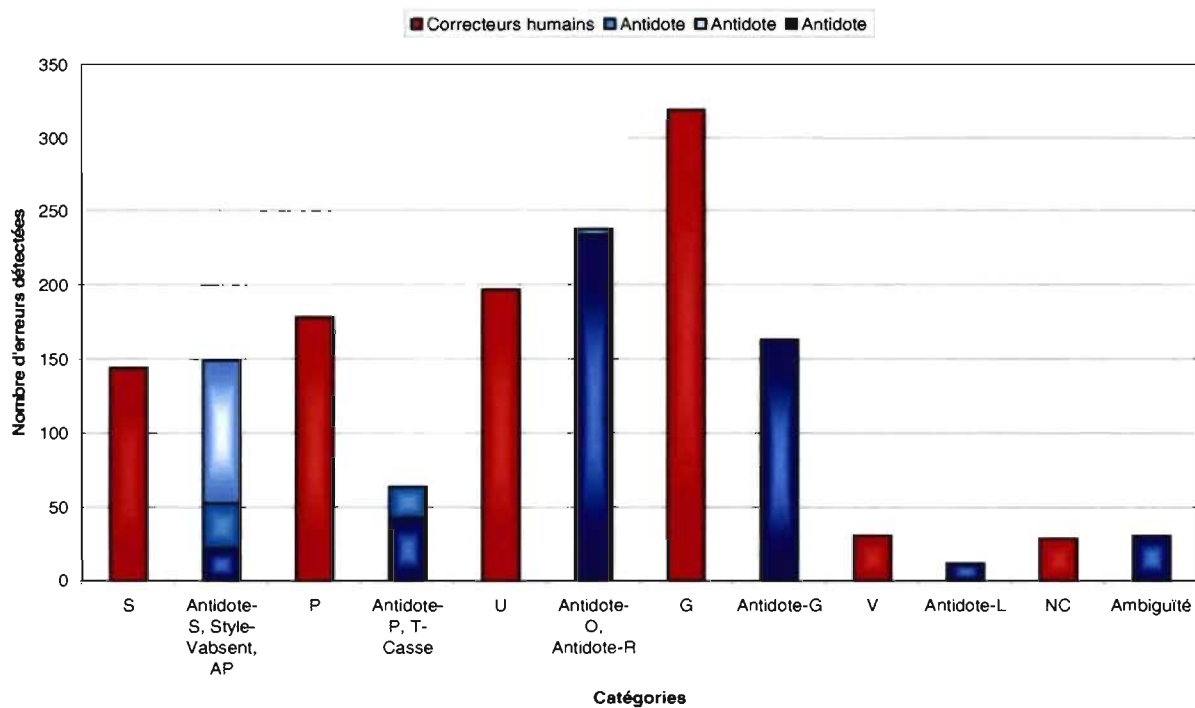


Fig. 7 - Nombre d'erreurs détectées selon les catégories – textes faibles

Tableau XI

Nombre d'erreurs détectées selon les catégories – textes moyens

Catégories		Type de correcteur	Nombre d'erreurs détectées			Total
Syntaxe	Syntaxe (S)	MELS	109			109
	Syntaxe (Antidote-S), Style-Vabsent, Analyse partielle (AP)	Antidote	11	34	43	88
Ponctuation	Ponctuation (P)	MELS	92			92
	Ponctuation (Antidote-P), Typographie (T-Casse)	Antidote	31	15		46
Orthographe d'usage	Orthographe d'usage (U)	MELS	73			73
	Orthographe d'usage (Antidote-O), Rédaction (Antidote-R)	Antidote	121	4		125
Orthographe grammaticale	Orthographe grammaticale (G)	MELS	197			197
	Orthographe grammaticale (Antidote-G)	Antidote	119			119
Vocabulaire	Vocabulaire (V)	MELS	24			24
	Lexique (Antidote-L)	Antidote	7			7
Non classées	Non classées (NC)	MELS	12			12
Ambiguïté	Ambiguïté	Antidote	22			22

Comparaison du nombre d'erreurs détectées selon les catégories - textes moyens

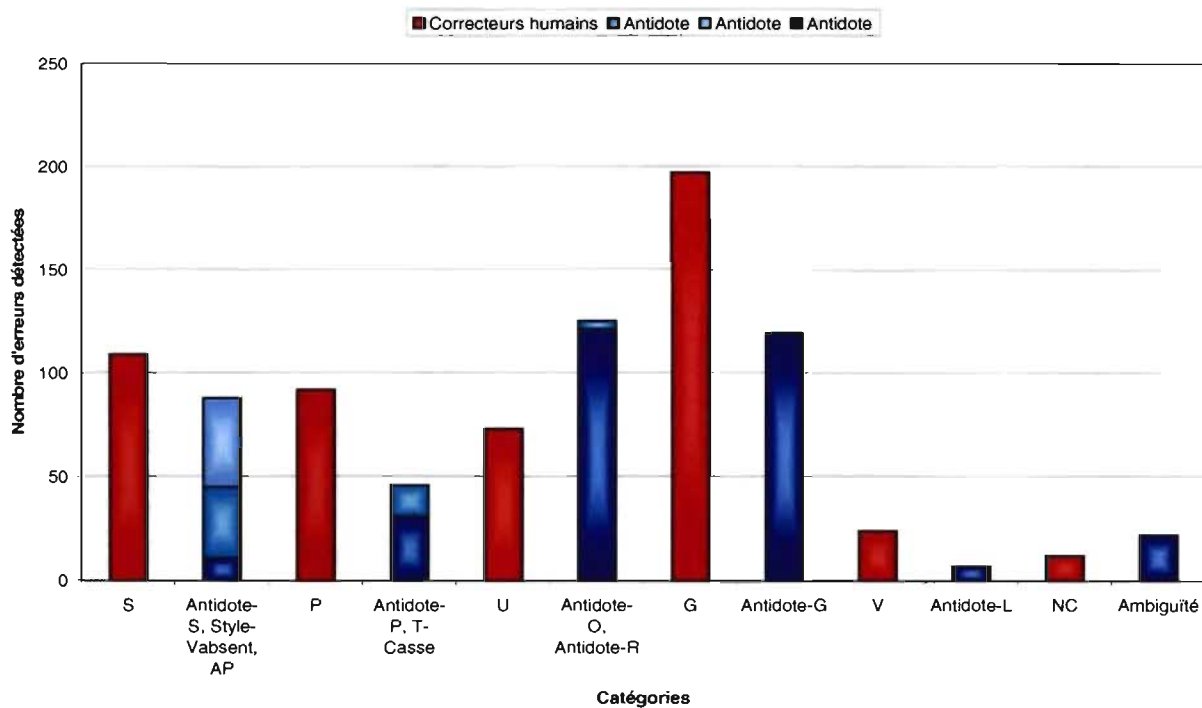


Fig. 8 - Nombre d'erreurs détectées selon les catégories – textes moyens

Tableau XII

Nombre d'erreurs selon les catégories – textes forts

Catégories		Type de correcteur	Nombre d'erreurs détectées			Total
Syntaxe	Syntaxe (S)	MELS	59			59
	Syntaxe (Antidote-S), Style-Vabsent, Analyse partielle (AP)	Antidote	24	21	30	75
Ponctuation	Ponctuation (P)	MELS	61			61
	Ponctuation (Antidote-P), Typographie (T-Casse)	Antidote	34	15		49
Orthographe d'usage	Orthographe d'usage (U)	MELS	52			52
	Orthographe d'usage (Antidote-O), Rédaction (Antidote-R)	Antidote	69	3		72
Orthographe grammaticale	Orthographe grammaticale (G)	MELS	109			109
	Orthographe grammaticale (Antidote-G)	Antidote	83			83
Vocabulaire	Vocabulaire (V)	MELS	16			16
	Lexique (Antidote-L)	Antidote	16			16
Non classées	Non classées (NC)	MELS	7			7
Ambiguïté	Ambiguïté	Antidote	26			26

Comparaison du nombre d'erreurs détectées selon les catégories - textes forts

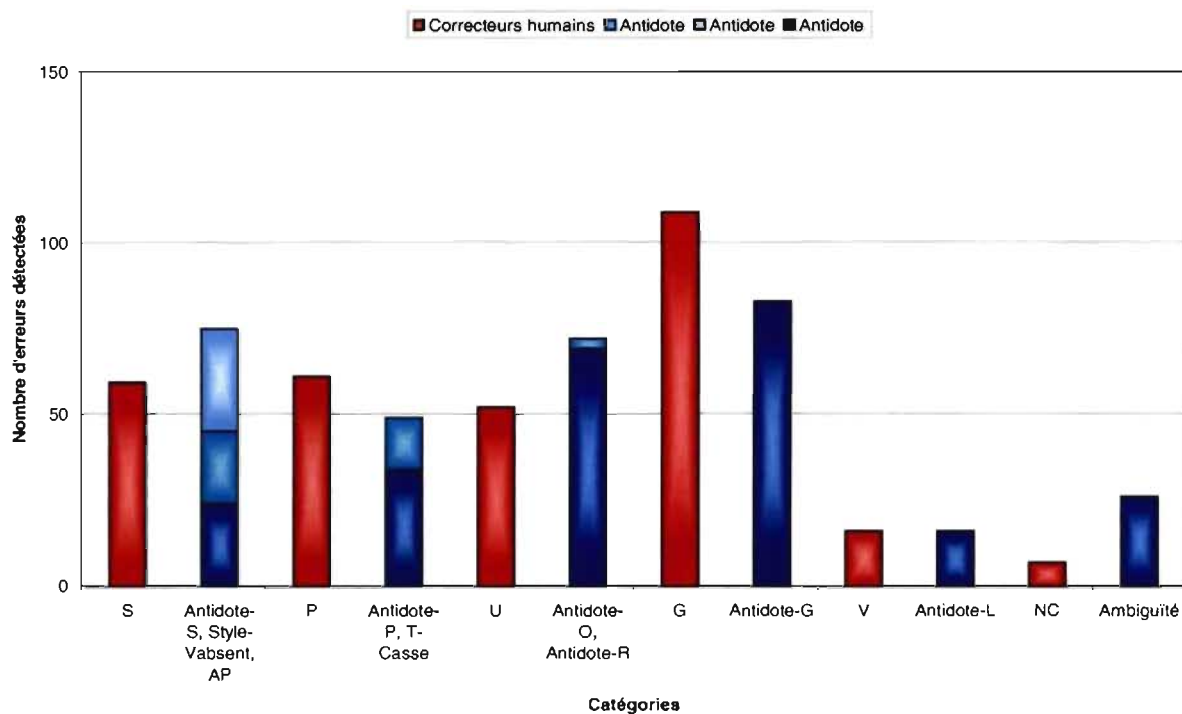


Fig. 9 - Nombre d'erreurs détectées selon les catégories – textes forts

Finalement, le dernier élément que nous avons observé pour l'analyse quantitative des données est le niveau de performance attribué à chacun des textes par les correcteurs. Les correcteurs du MELS ont classé les textes selon trois niveaux : *Faible*, *Moyen* ou *Fort*. Le classement d'*Antidote*, quant à lui, comporte 5 niveaux : *Novice*, *Eubage*, *Ovate*, *Maître* et *Druide*. Contrairement aux correcteurs humains, *Antidote* place chacun des textes sur une échelle de performance. Un texte peut donc se retrouver entre deux niveaux¹⁸. Pour notre analyse, nous avons dû attribuer un niveau précis à chacun des textes puisque la comparaison entre un niveau fixe et un niveau situé sur une échelle non métrique s'avère impossible. Le tableau ci-dessous (*Tableau XIII*) regroupe nos résultats.

Tableau XIII

Comparaison du niveau de performance attribué par les correcteurs

		Niveau de performance selon <i>Antidote</i>				
		<i>Novice</i>	<i>Eubage</i>	<i>Ovate</i>	<i>Maître</i>	<i>Druide</i>
Niveau selon le MELS	<i>Faible</i>	0	5	10	0	0
	<i>Moyen</i>	0	0	10	5	0
	<i>Fort</i>	0	0	2	10	3

On peut tout d'abord observer qu'aucun texte ne s'est retrouvé dans le niveau *Novice* d'*Antidote*. Les seuls textes se retrouvant dans le niveau *Eubage* sont des textes faibles. La situation est la même avec le niveau *Druide* : seuls des textes forts s'y trouvent. Le niveau *Ovate*, quant à lui, regroupe des textes des trois niveaux du MELS, soit des textes faibles, moyens et forts. De plus, on y trouve autant de textes faibles que de textes moyens. Pour le niveau *Maître*, on retrouve des textes forts et des textes moyens dans une proportion de 2 pour 1. Le niveau de performance attribué par *Antidote* et par les correcteurs humains n'est certes pas uniforme, mais on peut observer, d'une manière générale, une croissance similaire : plus le niveau attribué par les correcteurs humains augmente, plus le niveau de performance général attribué par *Antidote* augmente. Nous reviendrons, lors de l'interprétation des résultats (chapitre 5), sur la possible raison des différences observées dans l'attribution du niveau de performance et sur l'utilité ou non de cette échelle.

¹⁸ À titre d'exemple, voir la figure 3, page 40, où le texte se retrouve entre les niveaux *Ovate* et *Maître*.

4.2 Analyse qualitative

La seconde partie de notre analyse consistait à observer les différences dans l'identification et dans la classification des erreurs pour chaque type de correcteur. Notre attention s'est tout d'abord portée sur la correction effectuée par *Antidote*. Nous avons regardé en profondeur les situations dans lesquelles il « oubliait » des erreurs, les classait incorrectement, émettait des alertes, signalait des ambiguïtés ou était incapable d'effectuer l'analyse. Nous avons ensuite observé les failles dans la correction des correcteurs humains. Enfin, nous avons observé les différences de classement des erreurs entre les correcteurs du MELS, qui appliquent les directives du guide de correction, et *Antidote*, qui dépend du prisme de correction.

4.2.1 La correction d'*Antidote*

Il y a tout d'abord plusieurs situations différentes dans lesquelles *Antidote* peut omettre de signaler une erreur. La première relève de la grammaire textuelle et concerne les chaînes de reprise pour un référent donné. *Antidote* ne signale pas d'erreur de désignation du référent (ou erreur de reprise) dans l'extrait suivant : *Nous avons entendu parler du réchauffement de la planète. *Depuis quelques temps, ce dernier souffre [...]. Antidote* ne voit pas l'erreur soit parce qu'il n'arrive pas à retracer le référent ou parce qu'il identifiait le référent comme « réchauffement de la planète » et non « la planète ». *Antidote* peut également omettre de signaler une erreur lorsqu'un pronom n'a pas de référent. Dans un des textes, on retrouvait la phrase *Ils prévoient que les terres et les maisons de 118 millions de personnes risquent d'être inondés* dans un contexte où il était impossible d'attribuer une référence au pronom *ils*. Le référent de *ils* n'était pas mentionné précédemment. *Antidote* ne l'a pas souligné, peut-être parce qu'il a identifié un autre référent masculin pluriel ailleurs dans le paragraphe. *Antidote* a également de la difficulté à identifier plusieurs erreurs de syntaxe, dont l'ellipse du sujet. Les correcteurs humains avaient identifié une erreur dans la phrase **L'hiver est plus doux, passe vite*, tandis qu'*Antidote* n'a rien

signalé. Il n'identifie pas non plus les erreurs de syntaxe dans l'extrait **pareil que celle de New York et du gaspé*. Les erreurs concernant le temps de certains verbes passent également inaperçues, comme dans la phrase **Que serions-nous dans 20 ans?*, phrase qui, selon le contexte, aurait dû être écrite au futur simple de l'indicatif (*Que serons-nous dans 20 ans?*). *Antidote* peut également ne pas signaler une erreur dans le choix de la préposition, comme dans la phrase **La santé de la population sera affectée avec cette chaleur [...]*. Outre ces erreurs de syntaxe, *Antidote* omet également de signaler certaines erreurs de ponctuation. Il ne signale pas certaines virgules en trop, comme dans la phrase **Ce qui fait, que les rayons du soleil, se reflètent*. Il n'identifie pas non plus d'erreur lorsqu'une majuscule en trop apparaît au milieu d'une phrase (ex. : **Le Saumon riche en oméga-3 [...]*). Ceci s'explique peut-être par une confusion entre certains noms propres et noms communs. *Antidote* ne détecte pas non plus l'absence de point à la fin d'une phrase. L'absence de certaines virgules passe aussi inaperçue, comme dans l'extrait **nous les êtres humains, [...]* où il manque une virgule après le pronom. L'absence de certaines virgules pour détacher un organisateur textuel en début de phrase n'est pas toujours détectée non plus. Ainsi, *Antidote* ne signale pas d'erreur de ponctuation pour la phrase **En second lieu quelques-un seront heureux de cette situation [...]*. Le correcticiel est toutefois capable de signaler la virgule manquante lorsque la phrase commence par « Alors ».

Dans certaines situations, *Antidote* peut signaler qu'il y a une erreur, mais la classer incorrectement. Ceci se produit, par exemple, lorsque le mot erroné pourrait être un nom propre, comme dans **feux de forest*. *Antidote* signale une erreur à « forest », mais il la classe comme une erreur typographique plutôt qu'orthographique. Le correcticiel effectue également un classement erroné des erreurs lorsqu'il y a un point dans un titre (ex. : **Les changements climatiques.*). Plutôt que de souligner qu'il y a un point en trop, *Antidote* signale plutôt une absence de verbe. *Antidote* semble ainsi être programmé pour analyser chaque phrase comme étant une série de mots allant d'une majuscule à un point. Ceci se remarque aussi lorsqu'un point sépare deux phrases qui devraient plutôt être juxtaposées à l'aide d'une virgule. Dans l'extrait suivant, **Il y a plus de feux de forest qui sont hors de contrôle. Ce qui fait qu'il ce*

produit de plus en plus d'incendies, le point devrait être remplacé par une virgule. Toutefois, *Antidote* analyse les deux phrases séparément et indique plutôt qu'il manque un verbe dans la seconde. *Antidote* semble donc prendre pour acquis que les points sont placés au bon endroit, même si ce n'est souvent pas le cas, particulièrement dans des textes faibles. Ceci entraîne inévitablement des erreurs de correction de la part du correcticiel.

Antidote a la possibilité d'émettre des alertes lorsqu'il souligne un mot, mais qu'il n'est pas certain s'il s'agit bel et bien d'une erreur. Les alertes doivent donc être vérifiées par l'utilisateur du correcticiel. *Antidote* émet des alertes principalement lorsqu'un nom commun pourrait être utilisé comme un nom propre. C'est le cas des différents points cardinaux (nord, sud,...). Il émet également des alertes lorsqu'il pourrait manquer une virgule en début de phrase, pour détacher un complément. Il a parfois raison de signaler une erreur, comme dans l'exemple suivant : **Comme solution il y a le covoiturage [...]*. Il se trompe toutefois dans d'autres situations : *La saison végétative sera plus longue mais les récoltes plus difficiles [...]*. Dans cet extrait, *Antidote* a émis une alerte au mot *végétative* en indiquant qu'il manquait une virgule, ce qui n'est en fait pas le cas. *Antidote* indiquera aussi une alerte pour « si on » et pour la marque du pluriel pour certains noms, notamment ceux précédés d'un déterminant négatif, comme dans les extraits suivants : *comme nous n'avons pas d'abris, les animaux n'auront plus de banquises*. Il est à noter que les correcteurs humains, contrairement à *Antidote*, n'ont pas signalé d'erreur pour ces deux extraits.

Outre les alertes, *Antidote* peut également signaler des ambiguïtés. Contrairement aux alertes, les ambiguïtés ne sont pas comptabilisées par le correcticiel dans le nombre d'erreurs. Il s'agit de mots soulignés par le correcticiel demandant à l'utilisateur de révéifier cette partie du texte. Nous avons constaté qu'*Antidote* identifie des ambiguïtés lorsqu'une phrase commence par *Même*, lorsqu'il y a une répétition de mots (ex : **il est trop trop tard*), lorsque le sujet de la phrase est difficile à identifier, comme dans *nous les humains, qui n'apportent [...]*, ou encore lorsqu'il y a un doute quant au nom avec lequel un adjectif doit s'accorder (ex : *plus de blizzard et de*

tempêtes hivernaux). Une ambiguïté est également signalée lorsque des homonymes sont acceptables, en d'autres termes lorsqu'une autre orthographe serait possible, selon le sens de la phrase. C'est le cas, entre autres, dans l'extrait suivant : *un meilleur climat à gagner [...]*. *Antidote* indique une ambiguïté puisque nous aurions également pu voir « un meilleur climat a gagné... ». *Antidote* laisse le soin à l'utilisateur du correcticiel de juger par lui-même de l'orthographe à privilégier selon le sens désiré.

Dans d'autres situations, le correcticiel peut être incapable d'analyser une phrase ou une partie de phrase. Il indique alors qu'il a effectué une analyse partielle. C'est ce qui s'est produit avec les extraits suivants : **Dans ce texte on vous informera sur [...]*, **On doit essayer de diminuer [...]* et **Si chacun de nous ferions [...]*. L'analyse partielle concerne principalement les erreurs de syntaxe, qui semblent être le point faible de sa correction.

Mis à part les erreurs oubliées ou mal classées, les alertes, les ambiguïtés et les analyses partielles, *Antidote* peut également identifier une erreur où il n'y en a pas, ce que nous avons défini précédemment comme un « bruit » (voir la section 2.2.8). Ceci se produit principalement lorsque le sujet et le verbe sont séparés par d'autres groupes du nom qui sont d'un nombre différent du sujet, comme dans la phrase *Même les animaux, qui seront mélangés quand il fera chaud en hiver, croiront qu'on est le printemps et débalanceront leur cycle*. Dans cette phrase, *Antidote* a identifié une erreur à « débalanceront », indiquant que le verbe devait être au singulier. Le noyau du sujet se trouvant loin dans la phrase (*animaux*), il a probablement identifié « on » comme sujet du verbe, ce qui a entraîné une analyse erronée de la phrase. Nous avons également observé d'autres erreurs faussement identifiées, relevant par exemple de la confusion entre « peut-être » et « peut être ».

4.2.2 La correction par les correcteurs humains

Pour les correcteurs humains, notre attention s'est principalement portée sur les situations dans lesquelles ils oublient d'identifier certaines erreurs. Ceci s'est d'abord vu lorsqu'un adjectif était employé à la place d'un nom, soit avec le mot *pathogènes* (**La sécheresse empêchera certains pathogènes de se développer*). Des erreurs d'orthographe sont également passées inaperçues (**cardiovasculaire* au lieu de *cardiovasculaire*, **phériphérie* à la place de *périphérie*), ainsi que certains accords grammaticaux, comme **les être humains*. Des erreurs d'accent (ex. : **décés*) ou de majuscule (ex. : **virus du nil*) sont aussi oubliées à l'occasion. Les correcteurs humains peuvent oublier de souligner l'absence de certaines virgules (ex. : **Le problème c'est que [...]*) ou la présence de points dans les titres. Certains pléonasmes sont également passés sous silence (ex. : **il suffit tout simplement [...]*).

4.2.3 Retour sur les erreurs courantes

À la section 2.1.6 du présent mémoire, nous avons dressé une liste d'erreurs couramment observées dans les textes d'élèves. Nous avons donc tenté de repérer ces erreurs dans les textes retenus pour notre analyse afin de voir si *Antidote* et les correcteurs humains détectaient correctement ces erreurs. Voici, dans le tableau XIV (page suivante), un résumé de nos observations :

Tableau XIV

Observation des erreurs courantes dans les textes

Erreur courante	Exemples repérés dans les textes	Antidote		Correcteurs humains	
		Exemple repéré	Classement	Exemple repéré	Classement
Syntaxe					
Construction des phrases négatives	<i>*si nous faisons pas d'avantage attention [...]</i>	oui	Syntaxe - Négative	oui	Syntaxe
Construction des phrases interrogatives	<i>*Pourquoi est-ce que la terre change t-elle?</i>	oui	Analyse partielle	oui	Syntaxe
Groupe nominal					
Genre du nom	<i>*une petite effort</i>	oui	Grammaire - Nom	oui	Grammaire
Phrases complexes					
Choix du mode	<i>*Il se pourrait que la température moyenne à la surface augmentera de 1°C à 3,5°C.</i>	oui	Grammaire - Verbe	oui	Syntaxe
Complétives interrogatives indirectes	<i>*Au cours de ce texte tu va voir <u>quesqui</u> pourrait y arriver [...]</i>	oui	Orthographe - Divers	oui	Grammaire
Choix du pronom relatif	<i>*...les maladies dont le Québec n'a jamais eues auparavant.</i>	non		oui	syntaxe
Sujets GN complexes	<i>*Les rayons du soleil entre dans la surface [...]</i>	oui	Grammaire - Verbe	oui	Grammaire

Le tableau XIV nous permet d'observer qu'*Antidote* est capable de repérer la majorité des erreurs courantes. Seules les erreurs relevant du choix du pronom relatif dans des phrases complexes n'ont pas été identifiées par le correcticiel. *Antidote* identifie et classe correctement les erreurs concernant la construction de phrases négatives, le

genre du nom et les sujets du verbe qui sont des groupes nominaux complexes. Il identifie également les erreurs dans la construction de phrases interrogatives, mais signale une analyse partielle plutôt qu'une erreur de syntaxe. Pour le choix du mode dans des phrases complexes, *Antidote* classe l'erreur dans la catégorie *grammaire-verbe* tandis que les correcteurs humains la classent dans la catégorie *syntaxe*. Enfin, pour la complétive interrogative indirecte, *Antidote* signale une erreur touchant l'orthographe d'usage (**quesqui*), mais ne signale pas l'erreur de grammaire. Nous ne pouvons donc pas savoir, à partir de cet exemple, si *Antidote* est capable de détecter les erreurs dans les complétives interrogatives indirectes.

4.2.4 Différences dans le classement des erreurs selon le type de correcteur

Après avoir identifié les forces et les faiblesses des deux types de correcteur, nous avons voulu observer les différences¹⁹ dans la classification de certaines erreurs. Tout d'abord, nous avons constaté que les catégories *orthographe grammaticale* et *orthographe d'usage* ne regroupent pas les mêmes éléments chez les deux types de correcteur. Pour les correcteurs humains, les erreurs d'homonyme du type **il ce produit* ou **similaire a celui de [...]* font partie de la catégorie *orthographe grammaticale*, tandis qu'*Antidote* classe plutôt ces erreurs dans la catégorie *orthographe d'usage*. La situation est la même avec les erreurs de conjugaison, comme **plusieurs maladies apparaîteront*; pour *Antidote*, il s'agit d'une erreur d'orthographe d'usage et pour les correcteurs humains, d'une erreur d'orthographe grammaticale. D'autres erreurs comme **animals* ou **chaques* sont réparties de la même façon, soit erreurs d'orthographe grammaticale pour les correcteurs humains et erreurs d'orthographe d'usage pour *Antidote*. Certaines erreurs étant considérées comme des erreurs d'orthographe d'usage par les correcteurs humains sont toutefois classées autrement par *Antidote*. C'est le cas avec l'abréviation **ect*. Les correcteurs humains comptent une erreur d'orthographe d'usage, tandis qu'*Antidote* en compte

¹⁹ Les différences rapportées dans cette section ont été observées systématiquement. C'est pourquoi nous généralisons nos observations.

deux, soit une de rédaction et une de ponctuation. Les erreurs dans les noms propres, comme **Amérique*, sont classées dans l'orthographe d'usage pour les correcteurs humains et dans la typographie pour *Antidote*. L'absence d'une catégorie *typographie* dans le guide de correction amène donc les correcteurs humains à voir les erreurs typographiques comme des cas particuliers d'orthographe d'usage ou de ponctuation. D'autres différences peuvent toucher le mode de verbe erroné. Dans la phrase **Il se pourrait que la température moyenne à la surface augmentera de 1°C à 3,5°C.*, les correcteurs humains indiquent qu'il y a une erreur de syntaxe, tandis qu'*Antidote* indique qu'il y a une erreur d'orthographe grammaticale. Enfin, les deux types de correcteurs ne semblent parfois pas d'accord sur ce qui doit être identifié comme une erreur ou non. Par exemple, l'absence de virgule devant *et* dans une phrase comme *Les grandes métropoles peuvent changer la météo locale et même le climat régional* est considérée comme une erreur de ponctuation par *Antidote*, mais les correcteurs humains n'indiquent rien. La même chose se produit pour *et on*; *Antidote* indique une erreur de syntaxe, soulignant qu'on devrait plutôt écrire «et l'on», mais les correcteurs humains n'identifient aucune erreur.

Le dernier point sur lequel nous désirons nous pencher concerne les erreurs de vocabulaire. Lors de l'analyse quantitative des données, nous avons observé qu'*Antidote* et les correcteurs humains ont identifié le même nombre d'erreurs de vocabulaire dans les textes forts. Un examen plus attentif des données nous a cependant permis de constater que les erreurs détectées par *Antidote* ne sont pas les mêmes que celles repérées par les correcteurs humains. Ces derniers ont identifié, entre autres, les erreurs de vocabulaire suivantes : *Notre planète essaye de nous avertir pour qu'on stoppe [...]*, *La santé des êtres humains sera donc endommagée par [...]*. *Antidote* n'a signalé aucune erreur dans les deux exemples précédents. Inversement, certaines erreurs de lexique repérées par *Antidote* n'ont pas été signalées par les correcteurs humains. C'est le cas, entre autres, de **feux de forêt*, qui est une impropriété lexicale pour *Antidote*. Enfin, certaines erreurs sont identifiées par les deux types de correcteur, mais sont classées différemment : **augmentation*, par

exemple, est une erreur lexicale pour *Antidote*, mais une erreur d'orthographe d'usage pour les correcteurs humains.

Les différences entre les deux types de correcteur quant à l'identification et la classification des erreurs sont donc nombreuses. Dans le chapitre qui suit, nous tenterons de regrouper l'ensemble de nos résultats afin de les interpréter.

Cinquième chapitre

Interprétation et discussion

Les résultats de notre recherche nous ont permis de dresser un bilan global des forces et des faiblesses du correcticiel *Antidote* et de répondre, par le fait même, à notre objectif de recherche qui était de comparer la correction effectuée par des correcteurs humains à celle effectuée par le correcticiel *Antidote* lors d'une épreuve de français au secondaire. Tout d'abord, l'analyse quantitative faite à partir du nombre d'erreurs repérées dans les textes par chacun des types de correcteur nous a permis de constater qu'*Antidote* repère plus d'erreurs que les correcteurs humains dans les textes de niveau fort. *Antidote*, dans notre recherche, repère moins d'erreurs que les correcteurs humains dans les textes faibles ou moyens, particulièrement pour la catégorie *punctuation*. L'efficacité du correcticiel semble donc, de façon générale, plus élevée lorsque les textes renferment moins d'erreurs. Soulignons toutefois que notre analyse quantitative sur le nombre d'erreurs repérées incluait également les erreurs faussement identifiées par les correcteurs. En effet, certaines des erreurs comptabilisées n'en étaient pas réellement, ce qui a pu fausser un peu notre comparaison. Toutefois, la vue d'ensemble de cette recherche nous montre que le correcticiel semble moins habile à corriger les textes faibles plutôt que les textes forts. Est-ce que ceci signifie qu'on ne peut utiliser le correcticiel pour des textes trop faibles? Probablement pas. Malgré ses faiblesses, le correcticiel peut sans doute apporter une certaine aide aux correcteurs humains.

Tout d'abord, nous avons constaté, lors de la comparaison des niveaux de performance attribués aux textes par chaque type de correcteur, qu'il pouvait y avoir des différences majeures entre le classement d'*Antidote* et celui du MELS. Le classement d'*Antidote* est effectué à partir du nombre d'erreurs par phrase. Or, le nombre d'erreurs, pour *Antidote*, ne comprend pas les analyses partielles et les ambiguïtés. Les erreurs se retrouvant dans ces catégories peuvent être importantes

dans des textes d'élèves, particulièrement dans des textes faibles. Le niveau de performance attribué par *Antidote* aura donc tendance à être plus haut que celui attribué par le MELIS lorsqu'un texte comporte plusieurs phrases dans lesquelles des analyses partielles ou des ambiguïtés ont été signalées. Un autre élément que nous avons remarqué est qu'aucun texte d'élève ne s'est retrouvé dans la catégorie la plus faible (*Novice*). Il est probable que la difficulté pour *Antidote* d'analyser des textes faibles comportant plusieurs erreurs de syntaxe l'empêche de détecter toutes les erreurs et de classer dans la catégorie *Novice* des textes qui devraient pourtant s'y retrouver. Le jugement d'*Antidote* n'est toutefois pas complètement erroné. Comme nous l'avons vu, les textes se retrouvant dans la catégorie *Eubage* étaient tous des textes faibles, tandis que ceux se retrouvant dans la catégorie *Druide* étaient tous des textes forts. Cependant, les catégories *Ovate* ou *Maître* pouvaient correspondre à plusieurs niveaux de textes. Devrait-on pour autant s'abstenir d'utiliser les statistiques émises par *Antidote*? Peut-être pas. Il serait pertinent de les utiliser pour classer rapidement les textes faibles identifiés dans la catégorie *Eubage* par *Antidote*, et les textes forts identifiés, quant à eux, dans la catégorie *Druide*. Les textes présents dans les catégories *Ovate* et *Maître* demandent toutefois l'application du jugement de l'utilisateur, puisqu'il n'y a pas de correspondance exacte entre ces catégories et celles du MELIS.

Nous avons vu, lors de l'analyse qualitative des données, que les correcteurs humains n'ont pas nécessairement les mêmes faiblesses que le correcticiel. En effet, les correcteurs humains qui ont participé à notre étude ont tendance à oublier certaines erreurs d'orthographe, catégorie qui est, en contrepartie, le point fort d'*Antidote*. Il est facile pour un correcteur humain de passer par-dessus une erreur comme un accent absent ou une lettre manquante dans un mot, tandis qu'*Antidote* a démontré une constance dans la correction de ce genre d'erreur. Nous avons d'ailleurs vu, lors de la description du fonctionnement du correcticiel²⁰, que celui-ci est programmé à l'aide d'un dictionnaire contenant plus de 100 000 entrées. Lorsqu'il effectue la vérification lexicale dans un texte, le correcticiel souligne tous les mots ne

²⁰ Voir la section 2.3.2

correspondant pas aux formes qu'il connaît. Il ne peut donc pas oublier de signaler un mot dans lequel un accent est absent ou une lettre est manquante. Notre recherche vient ainsi confirmer la force majeure des correcticiels, qui est la correction orthographique des textes. *Antidote* et les correcteurs humains ont toutefois certaines faiblesses en commun, comme la ponctuation. Les correcteurs humains ont tendance à laisser passer, à l'occasion, des erreurs dans l'utilisation des virgules ou des majuscules. *Antidote* fait également des erreurs dans la correction de la ponctuation, mais ces erreurs ne relèvent pas d'« oubli » ou de correction moins sévère dans certains textes. En effet, comme nous l'avons expliqué précédemment²¹, les correcteurs humains ne sont pas infaillibles et il peut leur arriver, à cause de la fatigue ou de la lassitude, de laisser passer certaines erreurs. Ils peuvent également manifester de l'empathie envers certains élèves et laisser passer volontairement quelques erreurs parce qu'ils jugent que le reste du texte est excellent ou parce qu'ils veulent donner une chance à l'élève. Ces sentiments qui influencent les correcteurs humains ne peuvent se retrouver dans la correction par ordinateur. Une des forces principales du correcticiel vient donc de sa constance et de son absence d'empathie. Ceci ne l'a tout de même pas empêché, dans cette recherche, de laisser passer lui aussi certaines erreurs de ponctuation, mais ces « oublis » dans la correction se sont manifestés lorsque les phrases étaient mal construites. Cette faiblesse était également présente pour la catégorie *syntaxe*.

Nous avons constaté que le principal point faible d'*Antidote* était la détection des erreurs de syntaxe. *Antidote* est capable de reconnaître l'absence d'un verbe dans une phrase ou un emploi incorrect d'une forme négative, mais, lorsque la structure de la phrase est trop boiteuse, il est incapable de faire l'analyse. Il émet souvent des analyses partielles, des alertes ou signale des ambiguïtés lorsque la construction de la phrase ne lui permet pas de l'analyser correctement. Cette faiblesse d'*Antidote* vient probablement du fait que, pour lui, une phrase est constituée d'une série de mots compris entre une majuscule et un point. Or, ce n'est pas toujours le cas dans des textes d'élèves. Plusieurs élèves ont tendance, à certains endroits dans leur texte, à

²¹ Voir la section 1.3

mettre un point à la place d'une virgule. Le correcteur humain peut comprendre cette erreur et indiquer que l'élève a commis une erreur de ponctuation. *Antidote*, cependant, ne peut comprendre qu'un point devrait plutôt être une virgule et il effectue son analyse avec l'idée que les éléments se trouvant entre la majuscule et le point constituent une phrase. Ceci illustre la différence entre les définitions graphique et syntaxique de la phrase. Selon Riegel, une phrase graphique est une « suite de mots délimitée par une lettre majuscule initiale et par une ponctuation forte finale » (Riegel, 2004, p. 103). Une phrase syntaxique, quant à elle, correspond à un « assemblage de mots grammatical, c'est-à-dire conforme à des règles de construction » (Riegel, 2004, p. 104). *Antidote* fonctionne uniquement avec une définition graphique. Le correcteur humain, pour sa part, utilise la définition qu'il juge la plus appropriée, compte tenu de l'intention qu'il perçoit chez le scripteur à partir du texte. Il est donc évident qu'*Antidote* présente une faiblesse dans la détection précise des erreurs de syntaxe, mais il peut toutefois apporter une certaine aide aux correcteurs humains en leur signalant les problèmes potentiels, les phrases possiblement mal construites qu'il faut observer plus attentivement.

Enfin, le principal avantage à utiliser *Antidote* est son analyse beaucoup plus précise des erreurs détectées. Alors que les correcteurs humains classent les erreurs selon quatre catégories (cinq en incluant le vocabulaire), *Antidote* offre sept catégories de base, chacune se divisant encore en d'autres catégories plus précises, allant entre trois et huit. Les correcteurs humains n'ont pas le temps de noter, lorsqu'ils font la correction des textes, si l'erreur de grammaire qu'ils viennent d'indiquer relève d'un mauvais accord de l'adjectif, du participe passé, du verbe, etc. Or, *Antidote* peut, quant à lui, fournir cette information et même présenter un graphique à bandes illustrant, en pourcentage, la répartition des erreurs contenues dans le texte. On doit toutefois nuancer cet avantage. Comme nous l'avons déjà mentionné, *Antidote* a des faiblesses et il ne peut identifier avec précision toutes les erreurs contenues dans le texte. L'utilisateur du correcticiel ne peut donc se fier aux pourcentages présentés par *Antidote* dans son graphique à bandes, puisque, comme nous l'avons vu, plusieurs erreurs n'auront probablement pas été comptabilisées. De plus, certaines erreurs sont

parfois mal classées alors que d'autres erreurs identifiées par le correcticiel n'en sont pas réellement, ce qui rend encore une fois peu précise l'utilisation de la répartition des erreurs selon le pourcentage. Nous devons également nous rappeler que les correcteurs humains et *Antidote* ne s'entendent pas sur la classification de certaines erreurs. Il faudrait donc trouver un moyen d'uniformiser les catégories d'*Antidote* et celles du guide de correction du MELS si nous désirons utiliser les outils statistiques d'*Antidote*.

Sixième chapitre

Conclusion

Cette recherche avait pour objectif d'étudier les différences entre un correcticiel et un correcteur humain quant à leur correction respective de la qualité de la langue dans des textes d'élèves du secondaire. Nous désirions également juger de l'utilité d'un outil de correction comme *Antidote* lors des épreuves de français.

Nous pouvons conclure que le logiciel utilisé, *Antidote RX*, détecte moins d'erreurs que les correcteurs humains ayant participé à cette recherche dans les textes faibles de notre échantillon. Cette différence diminue toutefois lorsque le niveau du texte corrigé augmente pour atteindre un nombre d'erreurs détectées comparable aux correcteurs humains dans les textes forts. En d'autres mots, plus le niveau du texte est élevé, plus la correction d'*Antidote*, en terme de nombre d'erreurs détectées, se rapproche de la correction effectuée par les correcteurs humains du MELS. Nous avons également été en mesure d'identifier les points forts d'*Antidote*, soit la détection des erreurs d'orthographe d'usage plus constante que chez les correcteurs humains et l'analyse plus détaillée des erreurs, ainsi que ses points faibles, qui touchent principalement le domaine de la syntaxe. De nombreuses différences ont aussi été observées quant à la classification de certaines erreurs.

Notre étude est, bien évidemment, limitée. Elle est basée sur un corpus de 45 textes d'élèves de deuxième année du premier cycle du secondaire présentant des erreurs variées. Seuls des textes courants de type descriptif ont été soumis au correcticiel. Ainsi, nous ne pouvons généraliser nos résultats pour des textes courants relevant d'un autre type ou pour des textes littéraires ou encore pour des textes d'un autre niveau scolaire. Un texte de type narratif, par exemple, comporterait des temps de verbe différents (passé composé, passé simple, imparfait), ce qui mettrait *Antidote* à l'épreuve sur d'autres aspects de la langue. Nous ne pouvons savoir, à partir de notre

étude, si *Antidote* aurait le même niveau d'efficacité avec des textes de type différent. Nous ne pouvons également savoir si les correcteurs du MELS auraient détecté le même nombre d'erreurs si les textes n'avaient pas été corrigés une première fois par des enseignants. Il faut donc garder en tête que les résultats obtenus sont valides pour notre échantillon, mais qu'ils ne sont peut-être pas généralisables.

Finalement, un des objectifs de notre recherche consistait à juger de l'utilité d'un tel outil de correction lors des épreuves de français. Il est évident que les situations d'évaluation certificatives soumises par le MELS comportent des enjeux critiques ne laissant pas place à l'erreur lors de la correction. Un correcticiel comme *Antidote* n'a pas été développé sur mesure pour répondre aux besoins du MELS. Pourrait-il quand même être utilisé? Comme second correcteur pour vérifier certains éléments, comme les erreurs d'orthographe d'usage, peut-être. *Antidote* présente des forces que nous ne retrouvons pas chez les correcteurs humains ayant participé à cette recherche, comme la catégorisation plus détaillée des erreurs. Sa difficulté à analyser les erreurs de syntaxe rend toutefois son utilisation ardue pour la correction de textes d'élèves du secondaire. De plus, le classement des erreurs effectué par *Antidote* n'est pas toujours conforme au classement du MELS. Il est important, dans une évaluation certificative, de ne pas uniquement comptabiliser les erreurs correctement, mais de les classer adéquatement. En effet, un élève peut perdre tous ses points en syntaxe et ponctuation, mais réussir quand même l'examen à cause de sa force en orthographe, d'où l'importance de classer correctement les erreurs repérées dans les textes. Il serait intéressant de réaliser une étude semblable avec un correcticiel développé sur mesure pour les besoins du MELS. Nous pouvons, d'ici là, nous rappeler que les correcticiels ne sont pas nécessairement à écarter du revers de la main et qu'avec quelques modifications, ces outils pourraient véritablement être utilisés lors de la correction de textes d'élèves.

Bibliographie

Anctil, D. (2005). *Maîtrise du lexique chez les étudiants universitaires: typologie des problèmes lexicaux et analyse des stratégies de résolution de problèmes lexicaux*. Mémoire de Maîtrise, Université de Montréal, Montréal.

Bennett, R. E. (2006). Inexorable and Inevitable: The Continuing Story of Technology and Assessment. *Computer-Based Testing and the Internet: Issues and Advances*. D. Bartram and R. K. Hambleton. Chichester, John Wiley & Sons, Ltd.: 201-217.

Boivin, M.-C., Pinsonneault, R. et Philippe, M.-E. (2003). *Bien écrire, la grammaire revue au fil des textes littéraires* (2e ed.). Laval: Beauchemin.

Boudreault, P. (2004). La recherche quantitative. Dans T. Karsenti et L. Savoie-Zajc, *La recherche en éducation : étapes et approches*. Sherbrooke : CRP, p. 151-179

Burstein, J., Kukich, K., Wolff, S., Lu, C., Chodorow, M., Braden-Harder, L., et al. (1998). Automated scoring using a hybrid feature identification technique. In *Proceedings of the Annual Meeting of the Association of Computational Linguistics*. Montréal.

Burston, J. (1999). *Antidote 98*. CALICO [version électronique]. Récupéré le 31 mars 2008 de [https://www.calico.org/p-72-Antidote%2098%20\(11999\).html](https://www.calico.org/p-72-Antidote%2098%20(11999).html)

Charnet, C. et Panckhurst, R. (1998). Le correcteur grammatical: un auxiliaire efficace pour l'enseignement? Quelques éléments de réflexion. *Apprentissage des langues et systèmes d'information et de communication* [version électronique], Vol. 1, 103-114. Récupéré le 31 mars 2008 de <http://alsic.u-strabg.fr>.

Conseil des ministres de l'Éducation. (2003). *Les élèves et l'écriture : Contexte canadien*. Toronto: Programme d'indicateurs du rendement scolaire (PIRS 2002).

Cordier-Gauthier, C. et Dion, C. (2003). La correction et la révision de l'écrit en français langue seconde : médiation humaine, médiation informatique. *Apprentissage des langues et systèmes d'information et de communication* [version électronique], Vol. 6. 29-43. Récupéré le 31 mars 2008 de <http://alsic.u-strabg.fr>

Désilets, M. (1999). *L'utilisation pédagogique d'un logiciel correcteur au primaire et au secondaire*. Thèse de Doctorat, Université de Montréal, Montréal.

Dessus, P. et Lemaire, B. (2004). Assistance informatique à la correction des copies. In E. Gentaz & P. Dessus (Eds.). *Comprendre les apprentissages, Sciences cognitives et éducation* (pp. 205-220). Paris : Dunod.

De Viller, M.-É. (2003). *Multidictionnaire de la langue française* (4^e ed.). Montréal : Québec/Amérique.

Dikli, S. (2006). Automated Essay Scoring. *Turkish Online Journal of Distance Education-TOJDE*, 7(1), 49-62.

Druide (2006). *Antidote RX*. [version électronique]. Récupéré le 5 avril 2008 de <http://www.druide.com/telecharger/doc/posologie.pdf>

Druide (2007) *La rédaction administrative*. [cédérom] Antidote. Consulté le 25 juin 2009.

Druide (2008). *Antidote.info* [version électronique]. Récupéré le 2 février 2009 de http://www.druide.com/a_correcteur.html

Elliot, S. (2003). IntelliMetric™: From Here to Validity. In M. D. Shermis & J. C. Burstein (Eds.), *Automated Essay Scoring: A Cross-Disciplinary Perspective* (pp. 71-86). Mahwah: Lawrence Erlbaum Associates.

Foltz, P. , Laham, D. et Landauer, T. (1999). The Intelligent Essay Assessor: Applications to Educational Technology. *Interactive Multimedia Electronic Journal of Computer-Enhanced Learning*, 1(2). Winston-Salem, NC: Wake Forest University. Récupéré le 31 mars 2008 de <http://imej.wfu.edu/articles/1999/2/04/printver.asp>

Hutchison, D. (2007). An evaluation of computerized essay marking for national curriculum assessment in the UK for 11-years-old. *British Journal of Educational Technology*, 38(6), 977-989.

Jacquet-Pfau, C. (2001). Correcteurs orthographiques et grammaticaux: Quel(s) outil(s) pour quel rédacteur? *Revue française de linguistique appliquée*, Vol. VI 2001/2, 81-94.

Kukich, K. (2000). Beyond Automated Essay Scoring. *IEEE Intelligent Systems*, 15(5), 27-31.

Landauer, T. K., Laham, D. et Foltz, P. W. (2003). Automated Scoring and Anotation of Essays with the Intelligent Essay Assessor™. In M. D. Shermis & J. C. Burstein (Eds.), *Automated Essay Scoring: A Cross-Disciplinary Perspective* (pp. 87-112). Mahwah: Lawrence Erlbaum Associates, Inc.

Ministère de l'Éducation, du Loisir et du Sport (2008). *Guide de correction, prototype d'épreuve, FLE 2^e secondaire*. Gouvernement du Québec.

Ministère de l'Éducation du Québec (2006) *Programme de formation à l'école québécoise, enseignement secondaire, premier cycle*, Gouvernement du Québec.

Morin, R. (1995). *Sur l'intégration du correcticiel à la didactique du thème français*. Thèse de maîtrise, Université d'Ottawa, Ottawa.

Nadeau, M. et Fisher, C. (2006). *La grammaire nouvelle – La comprendre et l'enseigner*. Montréal : Gaëtan Morin éditeur.

Nichols, P. (2005). *Evidence for the interpretation and use of scores from an Automated Essay Scorer*: Pearson Educational Measurement.

Organisation de coopération et de développement économique. (2001). *Les nouvelles technologies à l'école: Apprendre à changer*. Paris: OCDE.

Page, E. B. (2003). Project Essay Grade: PEG. In M. D. Shermis & J. C. Burstein (Eds.), *Automated Essay Scoring: A Cross-Disciplinary Perspective* (pp. 43-54). Mahwah: Lawrence Erlbaum Associates.

Page, E. B. (1994). Computer Grading of Student Prose, Using Modern Concepts and Software. *Journal of Education Measurement*, 62(2), 127-142.

Powers, D.E., Burstein, J.C., Fowles, M.E. et Kukich, K. (2001). *Stumping e-rater: Challenging the validity of automated essay scoring*. (GRE Board Research Report No. 98-08bP). Princeton, NJ: Educational Testing Service. Récupéré le 2 février 2009 de <http://www.ets.org/Media/Research/pdf/RR-01-03-Powers.pdf>

Riegel, M., Pellat, J.-C. et Rioul, R. (2004). *Grammaire méthodique du français*. Paris: Quadrigue / PUF.

Rudner, L. et Gagne, P. (2001). An overview of three approaches to scoring written essays by computer. *Eric Digest* [version électronique]. Récupéré le 4 avril 2008 de http://eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/19/5e/46.pdf

Shermis, M. D., Mzumara, H. R., Olson, J. et Harrington, S. (2001). On-line Grading of Student Essays: PEG goes on the World Wide Web. *Assessment & Evaluation in Higher Education*, 26(3), 247-259.

Wang, J. et Brown, M.S. (2007). Automated Essay Scoring Versus Human Scoring : A Comparative Study. *Journal of Technology, Learning, and Assessment*, 6(2). Récupéré le 2 février 2009 de <http://www.jtla.org>.

Wolfe, E. W. (1997). The relationship between reading style and scoring proficiency in a psychometric scoring system. *Assessing Writing*, 4, 83-106.

Yang, Y., Buckendahl, C.W., Juskiewicz, P. J. et Bholá, D.S. (2002). A review of strategies for validating computer automated scoring. *Applied Measurement in Education*, 15(4), 391-412.

Annexe 1

Grille de correction du MELS

GRILLE D'ÉVALUATION – Échelle descriptive Nom de l'élève : _____ Groupe : _____
1^e cycle du secondaire (hypothèse 1)

Écrire des textes variés – Informer en élaborant des descriptions

Critère	Échelon 1 Manifestation d'une compétence très peu développée	Échelon 2 Manifestation d'une compétence peu développée	Échelon 3 Manifestation d'une compétence acceptable	Échelon 4 Manifestation d'une compétence assurée	Échelon 5 Manifestation d'une compétence marquée
1. Adaptation à la situation de communication	Présente peu d'éléments pertinents. ou Présente la plupart des informations sans les formuler dans ses mots.	Présente des descriptions liées au sujet traité. Tient compte de certains éléments de la tâche.	Élabore des descriptions la plupart du temps adaptées à la tâche.	Élabore des descriptions généralement pertinentes et les développe de façon appropriée. Suscite l'intérêt du destinataire.	Élabore des descriptions pertinentes et les développe de façon personnalisée. Suscite l'intérêt du destinataire par des moyens variés.
2. Cohérence du texte	Présente ses propos sans organisation évidente. ou Cumule des informations.	Présente ses propos de façon parfois discontinuée ou incomplète.	Organise son texte et fait progresser ses propos selon un ordre logique, malgré quelques maladresses.	Organise son texte de façon appropriée et fait progresser ses propos en établissant des liens logiques entre eux.	Organise son texte de façon appropriée et fait progresser habilement ses propos.
3. Utilisation d'un vocabulaire	Utilise des expressions ou des mots qui relèvent souvent de la langue orale ou familière.	Utilise des expressions ou des mots usuels dont plusieurs sont incorrects ou répétés.	Utilise des expressions ou des mots généralement appropriés, malgré des imprécisions ou des répétitions.	Utilise la plupart du temps des expressions et des mots précis et appropriés.	Utilise des expressions et des mots précis, variés et appropriés.
4. Construction de phrases et	Construit et ponctue correctement quelques phrases. (14-17 + erreurs)	Construit et ponctue de façon généralement correcte ses phrases de structure simple, mais celles de structure complexe comportent plusieurs erreurs. (9-16 erreurs)	Construit et ponctue de façon généralement correcte ses phrases de structure simple, mais celles de structure complexe comportent certaines maladresses. (6-12 erreurs)	Construit et ponctue ses phrases de façon généralement correcte, plusieurs présentant une certaine complexité. (3-8 erreurs)	Construit et ponctue correctement ses phrases, plusieurs présentant une certaine complexité, malgré de rares maladresses. (0-5 erreurs)
5. Respect des normes relatives	Orthographe son texte en respectant peu les normes relatives à l'orthographe. (26-32 + erreurs)	Orthographe son texte de façon peu assurée. Effectue parfois les accords simples et orthographe correctement certains mots courants. (17-31 erreurs)	Orthographe son texte de façon acceptable. Effectue plusieurs accords simples et orthographe de façon généralement correcte les mots courants. (12-20 erreurs)	Orthographe son texte de façon satisfaisante. Effectue plusieurs accords et orthographe de façon correcte les mots courants. (7-16 erreurs)	Orthographe son texte de façon très satisfaisante. Effectue la plupart des accords et orthographe les mots de façon généralement correcte, plusieurs étant moins courants. (0-8 erreurs)

Appréciation globale

Note. – Pour se situer à l'échelon 3 ou plus, l'élève doit avoir au moins une cote 3 aux critères 4 et 5.

Échelon 1 Manifestation d'une compétence très peu développée	Échelon 2 Manifestation d'une compétence peu développée	Échelon 3 Manifestation d'une compétence acceptable	Échelon 4 Manifestation d'une compétence assurée	Échelon 5 Manifestation d'une compétence marquée
Écrit un texte sommaire ou peu adapté à la tâche dans lequel la langue est très peu maîtrisée.	Écrit un texte qui traite sommairement du sujet et dans lequel la langue est peu assurée.	Écrit un texte adapté à la tâche et dans lequel la langue est d'un niveau acceptable.	Écrit un texte dont le contenu est développé de façon appropriée et dans lequel la langue est d'un niveau acceptable à satisfaisant.	Écrit un texte dont le contenu est développé de façon personnalisée et dans lequel la langue est d'un niveau satisfaisant à très satisfaisant.

GRILLE D'ÉVALUATION – Échelle descriptive Nom de l'élève : _____ Groupe : _____
1^e cycle du secondaire **(hypothèse 2)**

<i>Écrire des textes variés – Informer en élaborant des descriptions</i>					
Critère	Échelon 1 Manifestation d'une compétence très peu développée	Échelon 2 Manifestation d'une compétence peu développée	Échelon 3 Manifestation d'une compétence acceptable	Échelon 4 Manifestation d'une compétence assurée	Échelon 5 Manifestation d'une compétence marquée
1. Adaptation à la situation de communication	Présente peu d'éléments pertinents. ou Présente la plupart des informations sans les formuler dans ses mots.	Présente des descriptions liées au sujet traité. Tient compte de certains éléments de la tâche.	Élabore des descriptions la plupart du temps adaptées à la tâche.	Élabore des descriptions généralement pertinentes et les développe de façon appropriée. Suscite l'intérêt du destinataire.	Élabore des descriptions pertinentes et les développe de façon personnalisée. Suscite l'intérêt du destinataire par des moyens variés.
2. Cohérence du texte	Présente ses propos sans organisation évidente. ou Cumule des informations.	Présente ses propos de façon parfois discontinue ou incomplète.	Organise son texte et fait progresser ses propos selon un ordre logique, malgré quelques maladresses.	Organise son texte de façon appropriée et fait progresser ses propos en établissant des liens logiques entre eux.	Organise son texte de façon appropriée et fait progresser habilement ses propos.
3. Utilisation d'un vocabulaire	Utilise des expressions ou des mots qui relèvent souvent de la langue orale ou familière.	Utilise des expressions ou des mots usuels dont plusieurs sont incorrects ou répétés.	Utilise des expressions ou des mots généralement appropriés, malgré des imprécisions ou des répétitions.	Utilise la plupart du temps des expressions et des mots précis et appropriés.	Utilise des expressions et des mots précis, variés et appropriés.
4. Construction de phrases et	Construit et ponctue correctement quelques phrases. (14-17 + erreurs)	Construit et ponctue de façon généralement correcte ses phrases de structure simple, mais celles de structure complexe comportent plusieurs erreurs. (9-16 erreurs)	Construit et ponctue de façon généralement correcte ses phrases de structure simple, mais celles de structure complexe comportent certaines maladresses. (6-12 erreurs)	Construit et ponctue ses phrases de façon généralement correcte, plusieurs présentant une certaine complexité. (3-8 erreurs)	Construit et ponctue correctement ses phrases, plusieurs présentant une certaine complexité, malgré de rares maladresses. (0-5 erreurs)
5. Respect des normes relatives	Orthographe son texte en respectant peu les normes relatives à l'orthographe. (30-40 + erreurs)	Orthographe son texte de façon peu assurée. Effectue parfois les accords simples et orthographe correctement certains mots courants. (21-39 erreurs)	Orthographe son texte de façon acceptable. Effectue plusieurs accords simples et orthographe de façon généralement correcte les mots courants. (12-25 erreurs)	Orthographe son texte de façon satisfaisante. Effectue plusieurs accords et orthographe de façon correcte les mots courants. (7-16 erreurs)	Orthographe son texte de façon très satisfaisante. Effectue la plupart des accords et orthographe les mots de façon généralement correcte, plusieurs étant moins courants. (0-8 erreurs)

Appréciation globale

Note. – Pour se situer à l'échelon 3 ou plus, l'élève doit avoir au moins une cote 3 aux critères 4 et 5.

Échelon 1 Manifestation d'une compétence très peu développée	Échelon 2 Manifestation d'une compétence peu développée	Échelon 3 Manifestation d'une compétence acceptable	Échelon 4 Manifestation d'une compétence assurée	Échelon 5 Manifestation d'une compétence marquée
Écrit un texte sommaire ou peu adapté à la tâche dans lequel la langue est très peu maîtrisée.	Écrit un texte qui traite sommairement du sujet et dans lequel la langue est peu assurée.	Écrit un texte adapté à la tâche et dans lequel la langue est d'un niveau acceptable.	Écrit un texte dont le contenu est développé de façon appropriée et dans lequel la langue est d'un niveau acceptable à satisfaisant.	Écrit un texte dont le contenu est développé de façon personnalisée et dans lequel la langue est d'un niveau satisfaisant à très satisfaisant.

Annexe 2

Grille pour l'analyse des erreurs

Annexe 3

Autorisation de *Druide informatique inc.*

Bonjour Madame Mireault,

Vous nous écrivez:

> j'effectue présentement, dans le cadre de mon mémoire de maîtrise,
> une recherche sur l'application d'Antidote sur des textes d'élèves
> du secondaire. Je tenais à avoir votre permission afin d'insérer 3
> captures d'écran d'Antidote dans mon mémoire.
>

Cette précaution envers nos droits intellectuels vous honore. Nous vous accordons cette permission avec joie. Je dirais même plus: la nature de vos travaux nous intéresse. Si nous pouvons vous aider davantage, n'hésitez pas à nous écrire. Et lorsque vous aurez terminé, nous aimerions connaître les résultats de vos travaux. S'il vous était possible de nous transmettre une copie électronique de votre mémoire, nous vous en serions obligés.

Merci de votre intérêt pour Antidote,

Éric Brunelle
Président
Concepteur principal d'Antidote

_____		Druide informatique inc.
_____		http://www.druide.com
_____		[information retirée / information withdrawn]
_____		