Discourse and Writing/Rédactologie
Volume 35, 2025
http://journals.sfu.ca/dwr

Article

Évaluation exploratoire du module de reformulation d'Antidote : apports, limites et implications pour la formation à la révision professionnelle

Philippe Anckaert
Université libre de Bruxelles, centre de recherche Tradital

Nathalie Lemaire
Université libre de Bruxelles, centre de recherche Tradital

Résumé

Depuis le lancement de ChatGPT en 2022, les diverses applications d'écriture automatique reposant sur les grands modèles de langage affectent tout particulièrement la didactique de la rédaction, de la traduction et de la révision professionnelles. Face à l'opacité des mécanismes de reformulation proposés par ces technologies, il incombe aux enseignants-chercheurs[1] de (se) former à l'utilisation critique de ces outils. C'est dans cette optique que nous avons mené une étude exploratoire sur les performances du moteur de reformulation intégré en décembre 2024 au logiciel d'aide à la rédaction et à la révision Antidote (Druide informatique), largement utilisé dans notre enseignement de la révision unilingue. Cette étude vise à identifier les forces et les faiblesses dudit reformulateur afin de déterminer pour quelles tâches et dans quelles conditions il serait complémentaire des fonctionnalités plus anciennes offertes par Antidote.

Abstract

Since the launch of ChatGPT in 2022, the various automatic writing applications based on large language models have had a significant impact on the teaching of professional writing, translation and editing. Given the lack of transparency surrounding the rephrasing mechanisms offered by these technologies, it is incumbent upon academic staff to train in the critical use of these tools. With this in mind, we conducted an exploratory study on the performance of the rephrasing engine integrated in December 2024 into Antidote (Druide Informatique), a writing and editing

Discourse and Writing/Rédactologie
Volume 35, 2025
http://journals.sfu.ca/dwr

support software, which is widely used in our teaching of monolingual editing. This study aims to identify the strengths and weaknesses of this reformulation tool in order to determine which tasks and under what conditions would complement the older features offered by Antidote.

Remerciements

Nous tenons à exprimer notre reconnaissance à Damien Hansen pour ses orientations bibliographiques éclairées, qui ont contribué à affiner le cadre méthodologique de cette étude. Son expertise dans l'évaluation des grands modèles de langage a permis d'intégrer les spécificités des outils d'intelligence artificielle générative dans une approche critique et scientifiquement fondée.

1. Introduction

Le logiciel d'aide à la rédaction et à la révision Antidote développé par la société québécoise Druide informatique est intégré de longue date dans notre enseignement des techniques éditoriales au sein du Master en traduction éditoriale et littéraire de l'Université libre de Bruxelles. L'intérêt pédagogique de cet outil, reconnu comme « logiciel d'utilité scolaire » dans plusieurs pays francophones (Druide informatique, 2010), réside historiquement dans la richesse de son dictionnaire des cooccurrences et dans la précision informative de son module de correction En effet, lorsque ce module, conventionnellement désigné par « le Correcteur » à partir de ce point signale des erreurs ou lance des alertes, les signaux sont accompagnés de menus contextuels posant *a minima* un diagnostic (pour les alertes portant sur des segments non analysables par la machine) et renvoyant (pour la plupart des autres anomalies) à des dictionnaires et guides linguistiques aussi complets que nuancés. Ces informations précises et documentées permettent aux utilisateurs[1] de garder la main sur leur processus de rédaction et de révision de manière éclairée.

À l'automne 2024, à l'issue d'une phase de test *bêta* entamée en décembre 2023, la version infonuagique d'Antidote a été mise à jour pour intégrer un module de reformulation basé sur les grands modèles de langage (Druide informatique, 2023a, désigné par « le Reformulateur » à partir de ce point Il s'agit de la première application d'intelligence artificielle générative dans l'écosystème Antidote. Comme son nom l'indique, ce module propose aux utilisateurs des reformulations de phrases complètes, voire d'alinéas entiers. Aux dires de ses concepteurs, le Reformulateur procède d'une technologie propre, développée en toute indépendance pour

Discourse and Writing/Rédactologie
Volume 35, 2025
http://journals.sfu.ca/dwr

garder le total contrôle de la qualité de son fonctionnement (Druide informatique, 2023b ; Thivierge Côté, 2025), ce qui le distingue des autres applications bien connues d'IA générative linguistique. Le Reformulateur d'Antidote n'en partage pas moins avec ces autres produits des logiques de fonctionnement opaques qui ne permettent pas d'identifier les critères qui ont présidé aux reformulations proposées (Rudolph *et al.,* 2023). Dans l'expérience des utilisateurs, le Reformulateur rompt donc avec le caractère informatif – et *ipso facto* très formateur – des fonctionnalités plus anciennes d'Antidote.

L'arrivée du Reformulateur dans cette suite logicielle soulève une question pédagogique : comment intégrer cette nouvelle fonction, qui repose sur les mécanismes complexes et peu transparents typiques de l'intelligence artificielle générative, dans une approche de l'enseignement qui valorise la recherche et l'analyse critique d'information ? Le Reformulateur est-il, comme le présentent les concepteurs du produit, complémentaire des fonctionnalités du Correcteur qui lui préexistait (Druide informatique, 2023a ; 2023b ; Thivierge Côté, 2025) ? Si oui, pour quels types d'anomalies rédactionnelles en particulier ?

La littérature scientifique n'offrant, à notre connaissance, pas encore de recul par rapport à l'expérience utilisateur du Reformulateur d'Antidote, nous nous sommes proposé de tester cette fonctionnalité aussi rigoureusement que possible, afin d'en évaluer les apports et les limites, de manière à pouvoir adapter nos pratiques pédagogiques en conséquence. Nous partageons en effet les préoccupations de Frœliger (2025) :

> En tant que formateurs aux métiers de la traduction, nous avons besoin d'ouvrir grand nos yeux et nos oreilles pour percevoir où va le marché, pour nous y adapter et pour peser sur ces évolutions. Avec différentes sources d'information, qui constituent notre biotope. Car nous devons, et c'est une des difficultés que nous rencontrons, essayer de parvenir à une vision synthétique, dans un univers évolutif, qui mette en outre en lien la formation professionnelle et la recherche. (Frœliger, 2025, p. 152)

Vu l'essor fulgurant des outils d'IA linguistique, il nous semble urgent de les soumettre à une analyse scientifique indépendante, visant à identifier, dans l'esprit d'Anctil (2023, p. 72), des outils d'IA générative sûrs, éthiques et adaptés, de manière à tester, ensuite, les meilleures manières de les intégrer au curriculum des futurs professionnels du langage, aux pratiques pédagogiques de leurs formateurs ainsi qu'aux dispositifs d'évaluation des compétences rédactionnelle et traductive (voir à propos de cette dernière Anckaert, 2023). Soulignons au passage qu'en ce qui concerne les compétences du traducteur-réviseur, les outils d'IA linguistique

Discourse and Writing/Rédactologie
Volume 35, 2025
http://journals.sfu.ca/dwr

occupent déjà une place importante dans le référentiel de compétences du Master européen en traduction (Commission européenne, 2022, p. 9, compétences 15 et 16 en particulier).

2. Objectifs de l'étude

Cette étude, menée sur des textes produits par nos étudiants et basée sur la comparaison de trois modes de révision (humaine, Correcteur et Reformulateur), vise à :

- Identifier les forces et faiblesses actuelles du Reformulateur : dans quelle mesure, et pour quels types de problèmes rédactionnels relevés par les évaluateurs humains et le Correcteur, le Reformulateur se révèle-t-il (in)capable de proposer une reformulation pertinente, permettant au rédacteur ou au réviseur de résoudre la difficulté rencontrée ?
- Formuler de premières recommandations pédagogiques : quels enseignements peut-on tirer de l'observation des performances du Reformulateur sur notre corpus pour former les futurs professionnels du langage à une utilisation efficace, sinon optimale ?

3. Méthodologie

Dès la conception de notre dispositif d'évaluation des performances du Reformulateur à partir d'un corpus de textes d'étudiants (voir section 3.2), nous nous sommes trouvés aux prises avec un problème méthodologique de volatilité des données, inédit dans notre pratique d'enseignants-chercheurs non spécialistes de l'intelligence artificielle générative.

D'une part, comme dans les autres applications linguistiques bien connues d'IA générative, le nombre et le contenu des reformulations proposées pour un même passage varient d'une requête à l'autre en l'intervalle d'une seconde. D'autre part, et c'est la deuxième caractéristique des grands modèles de langage, les mises à jour de l'outil sont aussi fréquentes qu'imprévisibles (Druide, 2023a).

Cette instabilité rend particulièrement difficile toute évaluation scientifique rigoureuse de la qualité et de la pertinence des reformulations. Il s'agit là d'un enjeu méthodologique documenté dans la littérature récente sur les grands modèles de langage, dont nous avons dû prendre notre parti. Kreutzer *et al.* (2025), par exemple, insistent sur la nécessité de dépasser les évaluations strictement automatisées en appelant à des évaluations humaines nuancées et à l'élaboration de protocoles d'évaluation fine. C'est dans cette perspective que nous avons mené notre étude expérimentale, en assumant le caractère instantané et contextuel des données analysées.

Discourse and Writing/Rédactologie
Volume 35, 2025
http://journals.sfu.ca/dwr

3.1 Sujets d'expérience

L'étude a été menée auprès des quinze étudiants inscrits en 2024-25 à l'unité d'enseignement « Techniques éditoriales » de l'Université libre de Bruxelles. Le groupe se composait de quatre étudiants de deuxième année du Master en traduction éditoriale et littéraire, et de onze étudiants du Master en langues et lettres romanes, à finalités « Monde du livre et de l'édition » et « Écritures et littératures ». L'âge moyen des participants était de 23 ans.

Tous les sujets étaient des francophones, au sens qu'ils avaient effectué l'ensemble de leur parcours scolaire et universitaire en français dans différents pays (par ordre alphabétique : Belgique, Cameroun, France et Suisse). Leur participation s'inscrivait dans le cadre d'un cours obligatoire de leur finalité.

Le cours de techniques éditoriales vise à développer chez les étudiants des compétences avancées en révision unilingue, en combinant maitrise grammaticale, usage expert des outils linguistiques, et réflexion critique sur les pratiques rédactionnelles. Il intègre une initiation aux différentes étapes du processus éditorial, une sensibilisation aux enjeux de lisibilité et d'inclusivité, ainsi qu'une familiarisation avec les environnements technologiques actuels (correcticiels, outils collaboratifs, logiciels d'aide à la rédaction). L'approche adoptée encourage une posture réflexive face aux pratiques de révision, entre rigueur normative et adaptation aux attentes du donneur d'ordre.

Les étudiants en traduction avaient déjà été initiés aux fonctions de base du logiciel Antidote au cours de leur premier cycle. Les étudiants romanistes, quant à eux, ont été invités à se mettre à niveau à l'aide de tutoriels, afin de suivre sans difficulté le module de prise en main des fonctions avancées de l'outil, programmé en début de quadrimestre.

3.2 Corpus et conditions expérimentales

Le corpus analysé se compose des quinze productions écrites authentiques, réalisées par nos étudiants dans le cadre d'un exercice évalué de rédaction de quatrième de couverture en décembre 2024. Cette tâche s'inscrivait dans une séquence pédagogique consacrée à l'étude des caractéristiques génériques, discursives et stylistiques propres à ce type de texte, typiquement susceptible d'être produit dans le cadre des stages en maison d'édition.

Les consignes invitaient les étudiants à concevoir une quatrième de couverture originale pour un ouvrage appartenant à l'un des trois domaines suivants : fiction, sciences humaines ou vie pratique. Deux options étaient proposées pour la réalisation de la tâche. La première consistait à

Discourse and Writing/Rédactologie
Volume 35, 2025
http://journals.sfu.ca/dwr

rédiger la quatrième de couverture d'un ouvrage existant, en fournissant l'hyperlien menant vers la description officielle sur le site de l'éditeur, ainsi que, le cas échéant, la version originale de la quatrième de couverture si l'ouvrage n'était pas traduit en français. La seconde option invitait les étudiants à imaginer la quatrième de couverture d'un ouvrage fictif, en précisant les paramètres éditoriaux inventés (titre, auteur, public cible, éditeur, prix, nombre de pages, contenu, illustration de couverture, etc.). Une variante de cette seconde option, proposée aux étudiants en manque d'inspiration, consistait à modifier un paramètre d'un ouvrage réel — par exemple, en transposant l'intrigue dans un autre contexte spatio-temporel — et à adapter en conséquence la quatrième de couverture, désormais associée à cette œuvre fictive.

Cette séquence intervenant en fin de période, les étudiants, habitués au maniement d'Antidote, étaient censés s'être autorévisés à l'aide de cet outil avant de remettre leur production. Il s'agissait d'une tâche à réaliser en autonomie (« à domicile »), autrement dit hors du contrôle des enseignants en termes de temps d'exécution et de ressources réellement mobilisées. La possibilité leur était donc explicitement laissée de recourir à une ou plusieurs IA génératives de leur choix, incluant, ***de facto***, le Reformulateur d'Antidote, qui venait d'être intégré dans la suite logicielle et que nous n'avions, dès lors, pas encore eu l'occasion de tester de manière approfondie, ni par nous-mêmes, ni ***a fortiori*** en classe avec les étudiants.

Chaque production devait s'assortir d'un commentaire réflexif dans lequel les étudiants explicitaient leur démarche rédactionnelle, les choix effectués, les difficultés rencontrées, ainsi que leur éventuel recours à une intelligence artificielle générative. En cas d'utilisation d'un tel outil, il leur fallait documenter leur processus (requête initiale, réécriture, révision), annexer les différentes versions produites (avant et après intervention de l'application d'IA générative), et proposer une évaluation critique de l'outil mobilisé. L'analyse de ces données, mise en perspective avec les résultats de la présente étude, fera l'objet d'un développement ultérieur dans le cadre d'une autre publication. Nous nous bornerons ici à constater que moins de la moitié des sujets d'expérience (7 sur 15) ont déclaré avoir recouru à une application d'intelligence artificielle au cours du processus de rédaction. Parmi eux, moins de la moitié (3 sur 7) ont déclaré avoir utilisé la commande « Retoucher » du Reformulateur.

Le corpus totalise 2789 mots. La longueur des textes le constituant varie fortement d'un sujet d'expérience à l'autre : de 94 à 311 mots (voir Tableau 2). On constate une forte dispersion des valeurs autour de la moyenne de 185,9 mots par texte, qui se traduit par un écart-type particulièrement élevé ($\sigma = 64{,}6$).

3.3 Grille d'évaluation

Discourse and Writing/Rédactologie
Volume 35, 2025
http://journals.sfu.ca/dwr

Pour appliquer le protocole d'étude comparative décrit sous 3.5, nous avons eu recours à une grille d'évaluation correspondant à une typologie des problèmes rédactionnels dérivée d'une recherche antérieure (Lemaire, 2022). Cette grille n'a donc pas été conçue à l'origine pour comparer les performances de révision d'évaluateurs humains avec celles d'outils d'aide à la correction, mais elle est utilisée, sous une forme simplifiée, comme référentiel dans le cadre de l'unité d'enseignement de la révision professionnelle qui a fourni le corpus caractérisé plus haut.

Dans notre grille de référence, les anomalies rédactionnelles à détecter en (auto)révision sont organisées suivant trois grandes classes : grammaire, grammaire textuelle et efficacité communicationnelle.

Pour commencer sont catégorisées comme problèmes de nature grammaticale les entorses au code normatif écrit de la langue française. Les problèmes de nature grammaticale s'entendent au sens large. Ils se subdivisent en erreurs de typographie, de ponctuation, d'orthographe d'usage, d'orthographe grammaticale, de syntaxe et de lexique. Ces problèmes ne concernent que les niveaux subphrastique ou phrastique. La figure 1 détaille les types de problèmes de nature grammaticale.

Nature	**Qualité affectée**	**Norme enfreinte**	**Type de problème**
Grammaire	Correction	Typographie	majuscule/minuscule
			italique/romain
			autres signes
		Ponctuation	virgule
			point
			autres signes de ponctuation
		Orthographe d'usage	entité nommée
			coquille
			autre
		Grammaire *stricto sensu*	accord
			temps/mode verbal
			pléonasme grammatical
			autre
		Syntaxe	sélection de la préposition
			sélection de la classe d'objets
			sélection de la construction verbale
			anacoluthe
			autre problème de syntaxe
		Lexique	impropriété
			sous-détermination lexicale

Discourse and Writing/Rédactologie
Volume 35, 2025
http://journals.sfu.ca/dwr

			autre problème lexical

Figure 1. Typologie des problèmes grammaticaux (Lemaire, 2022, p. 147)

Les problèmes de grammaire textuelle, qui forment la deuxième classe, s'observent sur les plans micro- et macrostructurels : subphrastique, phrastique, interphrastique et transphrastique. La figure 2 présente ces problèmes. Il s'agit de défauts qui affaiblissent les relations entre éléments de l'unité textuelle considérée. Ces relations peuvent être de l'ordre de la signification (ou cohérence référentielle) : exactitude, complétude de l'information, logique du raisonnement, etc. ; de la cohérence structurelle : harmonisation de la forme ou de la formulation des éléments se situant sur un même plan ; ou de la cohésion (ou dynamique du discours) : continuité thématique, progression thématique, procédés de coréférence, etc.

Discourse and Writing/Rédactologie
Volume 35, 2025
http://journals.sfu.ca/dwr

Nature	**Qualité affectée**	**Norme enfreinte**	**Type de problème**
Grammaire textuelle	Cohérence référentielle	Exactitude référentielle	erreur factuelle
			erreur arithmétique
			autre inexactitude référentielle
		Logique	erreur logique
			passage non figurable/non-sens
		Univocité	ambiguïté lexicale
			ambiguïté syntaxique
			ambiguïté coréférentielle
		Explicitation	ellipse (terme de comparaison manquant ; autres)
			allusion, sous-entendu non identifiable
			présupposé non identifiable
			autre non-dit/autre incomplétude
		Prise en charge énonciative	sélection de marqueur énonciatif
			absence de marqueur énonciatif
	Cohérence structurelle	Organisation méso- et macrotextuelle	découpage en paragraphes
			autre
		Harmonisation	cohérence terminologique
			cohérence orthographique
			libellé et numérotation des titres et des éléments de liste
			organisateurs textuels, énumératifs
			rupture de point de vue, de mode d'exposition
			non-correspondance (titre/contenu)
			autres éléments du texte
	Cohésion (dynamique de l'information)	Continuité thématique	(non-)récurrence des termes
			sélection du procédé de coréférence
			défaut de contraste sémantique
			défaut de résonance
			autre défaut de continuité référentielle
		Connexité	parataxe
			présence d'un connecteur indésirable
			sélection du connecteur
		Progression	manque de progression
			excès de densité informationnelle
			autre défaut de progression

Discourse and Writing/Rédactologie
Volume 35, 2025
http://journals.sfu.ca/dwr

Nature	**Qualité affectée**	**Norme enfreinte**	**Type de problème**
		Étagement	problème de regroupement
			manque de parallélisme syntaxique
			manque de parallélisme sémantique
			manque de coordination/du marquage hiérarchique
			problème de focalisation
			problème de séquentialité
		Annonce	omission de l'annonce
			annonce trompeuse
			déviation

Figure 2. Typologie des problèmes de grammaire textuelle (Lemaire, 2022, p.149)

La troisième classe présentée à la figure 3 s'intitule « Problèmes d'efficacité communicationnelle » et recouvre les défauts de fluidité, d'idiomaticité et d'adaptation discursive en lien avec les contraintes d'appropriation des énoncés et de leur énonciation (Charolles, 1981).

Discourse and Writing/Rédactologie
Volume 35, 2025
http://journals.sfu.ca/dwr

Nature	Qualité affectée	Norme enfreinte	Type de problème
Efficacité communicationnelle	Fluidité	Progression logique de la phrase	écran linguistique
			complication syntaxique
			ordre illogique des constituants
			autre problème de progression logique
		Équilibre de la phrase	déséquilibre syntaxique
			autre problème d'équilibre
		Concision de la phrase	redondance sémantique
			dilution sémantique
			autre problème de concision
		Variété de la construction des phrases	répétition lexicale
			répétition syntaxique
			répétition phonique
			autre problème de variété
		Expressivité de la phrase	lexie terne
			métaphore discutable
			autre problème d'expressivité
	Idiomaticité	Prédictibilité des cooccurrents	dissonance expressive
			autre problème de manque de prédictibilité
			cliché
			autre problème d'excès de prédictibilité
	Adaptation discursive	Adaptation à la typologie textuelle	calque culturel
			trace d'oralité (inadapté à l'écrit)
			lexie/construction trop (peu) spécialisée
			lexie/construction trop soutenue/familière
			lexie/construction vieillie/néologique
			jugement de valeur déplacé
			autre problème d'adaptation à la typologie textuelle
		Adaptation au lecteur cible	pertinence
			autre problème d'adaptation au profil du lecteur
		Adaptation au guide de style du mandant et à l'éthos de l'énonciateur	feuille de style, charte graphique, terminologie du mandant
			politique du calque sémantique/syntaxique
			politique de l'emprunt
			politique du régionalisme
			politique de l'écriture inclusive
			politiquement correct (âgisme, sexisme, racisme…)
			autre problème d'adaptation au mandant

Figure 3. Typologie des problèmes d'efficacité communicationnelle (Lemaire, 2022 p. 152)

Discourse and Writing/Rédactologie
Volume 35, 2025
http://journals.sfu.ca/dwr

3.4 Hypothèses

Avant de procéder à une évaluation des performances du Reformulateur et de sa complémentarité déclarée avec le Correcteur (Druide informatique, 2023a ; 2023b ; Thivierge Côté, 2025), nos hypothèses initiales pouvaient se résumer ainsi : d'une part, nous supposions que l'usage combiné de ces deux outils permettrait une amélioration significative de la qualité linguistique et stylistique des textes produits ; d'autre part, nous envisagions que chacun de ces modules, bien que performant individuellement, atteindrait son plein potentiel dans une dynamique d'interaction, où les suggestions du Reformulateur viendraient affiner les corrections grammaticales et orthographiques proposées par le Correcteur. Ces hypothèses, encore théoriques à ce stade, nécessitaient une validation empirique rigoureuse pour en mesurer la portée réelle dans des contextes d'usage variés.

3.4.1 Complémentarité supposée du Correcteur et du Reformulateur dans la détection et la résolution des problèmes de nature grammaticale *lato sensu*

De nombreuses années d'expérience dans l'utilisation du Correcteur, nous ont amenés au constat que les anomalies grammaticales sont, dans l'ensemble, bien prises en charge. Ainsi que précisé plus haut, cet outil intervient à deux niveaux. D'une part, il fournit des corrections explicites accompagnées d'explications pour les erreurs détectées. D'autre part, il émet des alertes signalées par un surlignage et un diagnostic général (par exemple : « Rupture »), sans proposer de correction précise.

Nous formulons l'hypothèse que le Reformulateur, via la commande « Retoucher », pourrait générer des reformulations grammaticalement correctes pour les segments où le Correcteur se limite à une alerte (= HYP 1a). Dans un premier temps, il s'agit donc de comparer les suggestions des deux outils à celles issues d'une révision humaine, en quantifiant les cas de plus-value, de convergence et de moins-value (= silence ou erreur). Dans un second temps, nous nous intéressons à leur complémentarité : le Reformulateur apporte-t-il une plus-value dans les cas où le Correcteur reste silencieux ou fait une proposition inappropriée (= HYP 1b) ?

3.4.2 Complémentarité supposée du Correcteur et du Reformulateur dans la détection et la résolution des problèmes de grammaire du texte

D'après notre expérience d'utilisateurs, le traitement des problèmes de grammaire du texte par le Correcteur demeure partiel. En ce qui concerne la cohérence référentielle, son action se limite

Discourse and Writing/Rédactologie
Volume 35, 2025
http://journals.sfu.ca/dwr

essentiellement à la vérification de la graphie des noms propres présents dans son dictionnaire, ainsi qu'à la détection de certaines ambiguïtés syntaxiques. Pour ce qui est de la cohérence structurelle, il identifie principalement des problèmes d'harmonisation — qu'il s'agisse de l'orthographe (traditionnelle ou rectifiée), des formats de dates ou de la numérotation. Quant à la dynamique de l'information (Dupont *et al.,* 1994), elle n'est abordée que de manière indirecte, à travers les menus « Lisibilité » et « Tournures » du module d'inspection stylistique.

Nous formulons l'hypothèse que le Reformulateur pourrait aller plus loin, en détectant et en corrigeant d'autres types de ruptures de cohérence référentielle (notamment des erreurs logiques), ainsi que des défauts liés à la progression et à l'organisation de l'information (= HYP 2).

3.4.3 Complémentarité supposée du Correcteur et du Reformulateur dans la détection et la résolution des problèmes d'efficacité communicationnelle

À l'exception de quelques alertes liées au registre, accessibles via le sous-menu « Vocabulaire » du Correcteur stylistique, les enjeux d'efficacité communicationnelle échappent en grande partie aux capacités de la technologie sur laquelle repose le Correcteur. L'expérience vise à déterminer si la sensibilité du Reformulateur — notamment en matière de fluidité, d'idiomaticité et d'adaptation au contexte discursif — peut apporter des améliorations là où le Correcteur reste inactif (= HYP 3).

3.5 Protocole de recherche

L'étude s'est déroulée en trois étapes, chacune visant à comparer les performances de trois approches dans l'identification et le traitement des problèmes rédactionnels relevés dans le corpus d'écrits étudiants : (i) la révision humaine, (ii) le Correcteur, et (iii) la commande « Retoucher » du Reformulateur.

3.5.1 Établissement du repérage de référence par révision humaine

La première phase a consisté en une révision des textes par les deux enseignants-chercheurs, auteurs du présent article. Les évaluateurs ont travaillé de manière indépendante, tout en ayant recours, à leur discrétion, à des outils d'aide à la correction (correcteur automatique, grammaires de référence, ressources imprimées ou en ligne), à l'exclusion du Reformulateur. Aucun protocole d'assistance n'a été imposé, et les deux réviseurs n'ont pas nécessairement utilisé les mêmes ressources ni opéré suivant les mêmes modalités.

Discourse and Writing/Rédactologie
Volume 35, 2025
http://journals.sfu.ca/dwr

Les évaluateurs se sont délibérément donné pour ligne de conduite d'adopter une posture plutôt puriste, mais seules les erreurs (de toute nature) pour lesquelles il y avait un consensus interannotateur ont été retenues. Cette façon de procéder visait à neutraliser la subjectivité des évaluateurs, l'éventuelle tendance à l'hypercorrection ou les cas d'idiosyncrasie. En cas de doute ou de désaccord entre les évaluateurs, la détection n'a pas été prise en compte dans l'analyse quantitative. Par ailleurs, lorsqu'un même type d'erreur apparaissait à plusieurs reprises dans un texte, seule la première occurrence a été prise en compte.

En vue de l'analyse quantitative des données, les détections validées à ce stade de l'analyse du corpus ont été consignées dans une base de données sous tableur, conventionnellement désignée par « le Fichier général » à partir de ce point.

3.5.2 Analyse des détections du Correcteur

La deuxième phase de l'étude a consisté à comparer les détections issues de la révision humaine avec celles générées par le Correcteur. Afin de maximiser la sensibilité de l'outil, tous les paramètres ont été réglés à leur niveau d'alerte maximal. Le profil de compétence de l'auteur a été défini comme « moyen », afin d'éviter un abaissement du seuil d'exigence. Le paramètre de région linguistique a été fixé sur « français de France », avec activation des alertes sur les régionalismes, permettant une évaluation au cas par cas. Les modules « Correcteur de langue » et « Correcteur de style » ont été systématiquement activés.

Les résultats ont été encodés dans le Fichier général selon trois cas de figure :

- plus-value : détection pertinente par le Correcteur, non identifiée par les réviseurs humains ;
- convergence : détection partagée par le Correcteur et les réviseurs humains ;
- moins-value : détection erronée ou inappropriée (hypercorrection), susceptible d'induire en erreur un utilisateur non expert.

3.5.3 Analyse des propositions de la commande « Retoucher » du Reformulateur

La troisième phase de l'étude a porté sur l'analyse des reformulations générées par le Reformulateur via la commande locale « Retoucher ». Comme les autres commandes du Reformulateur, « Retoucher » propose entre une et trois reformulations par segment, avec une sélection par défaut actuellement opérée de manière aléatoire (Thivierge Côté, 2025). Cette commande a été systématiquement appliquée aux segments jugés insatisfaisants par la révision

Discourse and Writing/Rédactologie
Volume 35, 2025
http://journals.sfu.ca/dwr

humaine et/ou par le Correcteur — autrement dit, dans les cas de convergence du correcteur automatique avec la révision humaine ou de plus-value du correcteur automatique par rapport à la révision humaine.

Les autres fonctionnalités du Reformulateur (« Réécrire », « Épicéniser », « Modérer », « Raccourcir »), jugées trop prolifiques ou éloignées des objectifs de cette étude, ont été écartées et pourront faire l'objet d'analyses ultérieures.

L'encodage des résultats dans le Fichier général s'est fondé sur les critères suivants :

- plus-value (+) : reformulation jugée pertinente, apportant une amélioration au texte original ;
- neutre (0) : reformulation sans effet notable, n'améliorant ni ne dégradant le texte ;
- moins-value (–) : reformulation inappropriée ou erronée, relevant d'une hypercorrection.

4. Résultats

Le Tableau 1 présente par ordre décroissant de fréquence les types de problèmes relevés dans le corpus de textes d'étudiants (E1 à E15) par la révision humaine et/ou par le Correcteur, que les détections de ce dernier constituent une plus-value, un cas de convergence ou une moins-value. Ces résultats nous renseignent sur la nature des faiblesses qui persistent dans les écrits de nos étudiants en dernière année de formation, même quand ils ont accès aux outils informatisés d'aide à la rédaction et à la correction. Mais, dans le cadre de la présente étude, la fonction de ce relevé consiste à servir de point de référence pour évaluer les performances du Reformulateur.

Tableau 1. Nature des problèmes détectés par les réviseurs humains et/ou le Correcteur, par ordre décroissant de fréquence

Type de problème	**Occurrences**									**Total**
complication syntaxique	E1	E8	E9	E10	E13	E13	E13	E14	E15	9
anacoluthe	E1	E1	E1	E8	E13	E15	E15			7
impropriété	E5	E5	E6	E7	E10	E10	E14			7
répétition lexicale	E1	E8	E8	E9	E9	E13				6
sélection construction verbale	E3	E6	E8	E8	E9	E10				6
virgule	E6	E6	E8	E9	E14	E14				6
dissonance expressive	E4	E7	E15							3
ellipse	E8	E8	E13							3
erreur factuelle	E1	E8	E9							3
erreur logique	E5	E7	E14							3
grammaire	E4	E8	E14							3

Discourse and Writing/Rédactologie
Volume 35, 2025
http://journals.sfu.ca/dwr

Type de problème	Occurrences									Total
pb adaptation typologie textuelle	E3	E7	E11							3
pb lexical	E2	E6	E15							3
pb manque prédictibilité	E1	E2								3
présupposé non identifiable	E1	E3	E13							3
sélection procédé coréférence	E4	E5	E11							3
accord	E6	E14								2
déséquilibre syntaxique	E8	E9								2
écran linguistique	E1	E13								2
majuscule	E3	E13								2
non-sens	E6	E11								2
ordre illogique constituants	E3	E12								2
orthographe usage	E6	E7								2
passage non figurable	E10	E11								2
pb variété	E8	E8								2
point	E8	E10								2
redondance sémantique	E4	E8								2
sélection connecteur	E1	E7								2
sélection préposition	E8	E10								2
typographie	E6	E15								2
adaptation au lecteur	E12									1
adaptation typologie textuelle	E14									1
allusion non identifiable	E10									1
ambiguïté syntaxique	E6									1
annonce trompeuse	E6									1
construction trop familière	E11									1
coquille	E7									1
découpage en paragraphes	E14									1
défaut de résonance	E3									1
excès de densité informationnelle	E15									1
entité nommée	E14									1
harmonisation orthographique	E3									1
harmonisation, autre élément	E7									1
italique	E8									1
lexie terne	E1									1
manque parallélisme syntaxique	E13									1
pb concision	E3									1
pb séquentialité	E10									1
pléonasme grammatical	E6									1
redondance lexicale	E11									1

Discourse and Writing/Rédactologie
Volume 35, 2025
http://journals.sfu.ca/dwr

Type de problème	Occurrences									Total
répétition syntaxique	E8									1
rupture de point de vue	E14									1
sous-détermination lexicale	E14									1
trace d'oralité	E15									1

Après ce premier relevé, nous avons évalué le potentiel de suggestion de variantes du Reformulateur. Tout étant relatif, nous avons procédé par comparaison avec les deux autres modes de révision : la révision humaine et celle effectuée par le Correcteur.

Pour les trois modes, nous avons enregistré, par étudiant et pour tout le corpus, le nombre de segments de phrases ayant fait l'objet 1°) d'une observation de la part des réviseurs humains, 2°) d'une alerte de la part du Correcteur et 3°) d'une ou plusieurs propositions émanant du Reformulateur. Pour la commodité, nous avons qualifié les trois types d'intervention de «°détection°».

Les résultats de ces différents décomptes figurent dans le Tableau 2.

Tableau 2. Nombre de détections par mode de révision

A	B	C	D	E	F	G	H	I	J
	Mots	Dét. hum. abs.	Dét. hum. rel.	Dét. Cor. abs.	Dét. Cor. rel.	Dét. Ref. abs.	Dét. Ref. rel.	Reform. abs.	Reform. rel.
E1	94	11	11,70	9	9,57	1	1,06	3	3,19
E2	156	2	1,28	6	3,85	3	1,92	6	3,85
E3	141	9	6,38	8	5,67	4	2,84	6	4,26
E4	168	3	1,79	5	2,98	2	1,19	4	2,38
E5	169	4	2,37	4	2,37	3	1,78	6	3,55
E6	189	12	6,35	13	6,88	4	2,12	10	5,29
E7	311	8	2,57	12	3,86	5	1,61	8	2,57
E8	139	7	5,04	7	5,04	0	0,00	0	0,00
E9	209	13	6,22	15	7,18	1	0,48	2	0,96
E10	284	10	3,52	12	4,23	1	0,35	6	2,11
E11	169	6	3,55	6	3,55	0	0,00	0	0,00
E12	171	2	1,17	4	2,34	1	0,58	4	2,34
E13	164	10	6,10	8	4,88	2	1,22	3	1,83
E14	123	12	9,76	14	11,38	2	1,63	3	2,44
E15	302	8	2,65	12	3,97	6	1,99	7	2,32
Total	2789	**117**	s.o.	**135**	s.o.	**35**	s.o.	**68**	s.o.
Moyenne	185,93	7,80	4,70	9,00	5,18	2,33	1,25	4,53	2,47
Ecart-type	64,59	3,71	3,09	3,72	2,60	1,80	0,84	2,83	1,46

Dans la colonne A figurent les identifiants anonymisés des étudiants ; dans la colonne B, le nombre de mots comme indication de la longueur de leurs productions ; dans les colonnes C, E et

Discourse and Writing/Rédactologie
Volume 35, 2025
http://journals.sfu.ca/dwr

G le nombre de détections en valeur absolue, et dans les colonnes D, F et H le nombre de détections en valeur relative, exprimée en pourcentage du nombre de mots du texte concerné. Les colonnes situées à l'extrême droite du tableau renseignent le nombre de propositions faites par le Reformulateur pour chacun des textes, sous I en valeur absolue et sous J en valeur relative.

Du tableau 2, il ressort que, pour le corpus dans son ensemble, le Reformulateur totalise seulement 35 détections contre 135 effectuées par le Correcteur et 117 par le tandem de réviseurs humains.

Le faible taux de détection du Reformulateur (Dét. Ref. abs., moyenne = 2,33) ne peut s'expliquer par un haut degré de compétence rédactionnelle des sujets d'expérience. En effet, dans cinq cas sur neuf où le nombre de détections est inférieur à la moyenne de 2,33 (voir E1, E9, E10, E13 et E14), les valeurs enregistrées pour la détection humaine (Dét. hum.) et/ou la détection du Correcteur (Dét. Cor.) s'avèrent supérieures à la moyenne obtenue pour ces deux autres modes de révision (respectivement 7,8 et 9,0).

L'examen des valeurs relatives révèle que celles-ci varient de 1,17 à 11,7 pour les détections humaines (Δ= 10,53) et de 2,34 à 11,30 pour le Correcteur (Δ= 8,96), alors que le différentiel est beaucoup plus modeste pour le Reformulateur (Δ=2,84). Les variations importantes enregistrées pour les détections humaines et le Correcteur se traduisent par des écarts-types élevés (resp. 3,09 et 2,60). Il apparaît donc que l'échantillon de population se caractérise par une grande hétérogénéité du niveau de compétence rédactionnelle, à considérer globalement comme plutôt bas pour des étudiants en fin de cycle. En conséquence, il est exclu d'attribuer le faible taux de détection du Reformulateur à un degré élevé de qualité des productions étudiantes.

Il semble en revanche exister un lien entre le nombre de détections par texte et la longueur du texte. Dans sept cas sur neuf (voir E1, E4, E8, E11, E12, E13, E14), soit dans 78 % des cas où le nombre de détections par le Reformulateur (0, 1 ou 2) est inférieur à la moyenne (2,33), on constate que ces valeurs se rapportent à des textes courts, à savoir des écrits comportant un nombre de mots inférieur à la moyenne de 185,9. Ce constat semble indiquer que la capacité du Reformulateur à détecter des problèmes rédactionnels s'accroît à mesure que la quantité de texte à traiter est plus importante.

Quant au nombre de propositions de reformulation (colonnes I et J), au total 68 pour l'ensemble du corpus et en moyenne 4,53 par texte, ce nombre ne semble entretenir aucun lien avec le nombre de détections (colonnes G et H), ni avec la longueur des textes et encore moins avec leur qualité rédactionnelle telle que mise en lumière par les deux autres types de détections.

Notons que le nombre très élevé de détections enregistrées pour le Correcteur (n = 135) s'explique par le choix des réglages mentionné sous 3.5.2. En maximisant la sensibilité de l'outil,

Discourse and Writing/Rédactologie
Volume 35, 2025
http://journals.sfu.ca/dwr

on favorise inévitablement la production d'alertes inappropriées. C'est ce qui ressort de l'analyse qualitative, opérée selon le protocole d'expérience décrit sous 3.5.2 et 3.5.3, des détections effectuées par le Correcteur et des propositions d'amélioration faites par le Reformulateur.

Le Tableau 3 illustre la proportion de détections/propositions qualifiées de «°judicieuses°» (= «°plus-value°» + «°convergence°»/«°neutre°») par rapport au nombre total de détections/propositions effectuées par les deux outils. Il est frappant de constater que le Correcteur n'affiche que 32 détections judicieuses sur 135, soit 23,7 % contre 36,7 % pour le Reformulateur (25 détections judicieuses sur 68). Alors que ces pourcentages semblent rapprocher en quelque sorte les deux outils sur l'axe de l'hypercorrection, l'examen détaillé du tableau, ligne par ligne, révèle en réalité une grande disparité de situations. Hormis les trois cas (E1, E12 et E13) où la proportion de détections judicieuses est (quasiment) la même pour le Correcteur et le Reformulateur, cette proportion se révèle nettement supérieure tantôt pour le Reformulateur (E2, E5 et E7), tantôt pour le Correcteur (E3, E8 et E15). Notons au passage que dans le cas de E8, le Reformulateur ne détecte rien alors que le Correcteur émet une alerte à sept reprises et propose trois corrections judicieuses. Cette grande disparité des résultats est confirmée par le calcul des coefficients de corrélation, effectué pour la totalité de l'échantillon : pour le rapport de corrélation entre le total des détections du Correcteur et des propositions du Reformulateur (colonnes B et D), r = 0,26 ; pour le rapport entre les détections et les propositions judicieuses de l'un et de l'autre outil (colonnes C et E), r = 0,04. Ces valeurs extrêmement faibles tendent à indiquer qu'il n'existe pas de lien structurel entre les deux modules étudiés.

Discourse and Writing/Rédactologie
Volume 35, 2025
http://journals.sfu.ca/dwr

Tableau 3. Proportion de détections judicieuses effectuées par le Correcteur et le Reformulateur

A	B	C	D	E
	Dét. Cor. tout	Dét. Cor. judic.	Dét. Ref. tout	Dét. Ref. judic.
E1	9	3	3	1
E2	6	1	6	3
E3	8	4	6	1
E4	5	1	4	1
E5	4	0	6	2
E6	13	3	10	5
E7	12	0	8	3
E8	7	3	0	0
E9	15	6	2	2
E10	12	1	6	2
E11	6	0	0	0
E12	4	2	4	2
E13	8	5	3	2
E14	14	1	3	1
E15	12	2	7	0
Total	135	32	68	25
Moyenne	9,00	2,13	4,53	1,67
Ecart-type	3,72	1,85	2,83	1,35

Discourse and Writing/Rédactologie
Volume 35, 2025
http://journals.sfu.ca/dwr

Tableau 4. Bilan qualitatif des propositions de reformulation

	P1			P2			P3		
	Plus-value	Neutre	Moins-value	Plus-value	Neutre	Moins-value	Plus-value	Neutre	Moins-value
E1	x					x			x
E2.S1			x		x			x	
E2.S2			x						
E2.S3	x					x			
E3.S1			x						
E3.S2			x			x			
E3.S3			x			x			
E3.S4	x								
E4.S1			x						
E4.S2	x					x			x
E5.S1			x			x			x
E5.S2		x			x				
E5.S3			x						
E6.S1			x			x			x
E6.S2	x					x	x		
E6.S3	x			x			x		
E6.S4			x						
E7.S1		x			x				
E7.S2	x								
E7.S3			x						
E7.S4			x			x			x
E7.S5			x						
E8									
E9.S1		x							
E9.S2	x								
E10.S1			x			x		x	
E10.S2			x			x			x
E11									
E12.S1	x								
E12.S2			x		x				x
E13			x		x			x.	
E14.S1	x								
E14.S2			x			x			
E15.S1		x				x			
E15.S2			x						
E15.S3			x			x			
E15.S4						x			
E15.S5		x							
Total prop.	10	5	20	1	5	15	2	3	7
Total valeurs cumulées			35			21			12

Discourse and Writing/Rédactologie
Volume 35, 2025
http://journals.sfu.ca/dwr

Du Tableau 4, il ressort que :

- deux textes (E8 et E11) n'ont suscité aucune reformulation°;
- hormis ces deux cas, les textes ont fait l'objet de 1 à 5 détections (codées ici de S1 à S5, car elles ont trait à un segment de texte)°;
- chaque segment jugé susceptible d'amélioration a donné lieu à une, deux ou trois propositions de reformulation (représentées ici par P1, P2 et P3), qui ont été réparties en trois colonnes selon leur pertinence («°plus-value°», «°neutre°» «°moins-value°») ;
- les détections donnant lieu à une seule proposition sont majoritaires (n = 16), suivies de près par celles ayant généré trois propositions de reformulation (n = 13), à la différence des détections à deux propositions largement minoritaires (n = 6)°;
- la pertinence et la qualité des reformulations sont sujettes à une grande variabilité, caractérisée par une nette dominance des «°moins-values°» par rapport aux deux autres cas de figure : en valeurs absolues, 20 sur 35 (57 %) pour les propositions P1, 15 sur 21 (71 %) pour les P2 et 7 sur 12 (58 %) pour les P3°;
- en valeurs absolues, que ce soit pour P1, P2 ou P3, les propositions inappropriées se révèlent systématiquement plus nombreuses que les propositions judicieuses (= «°plus-values°» + «°neutres°») (52 sur 68, soit 76 %).

Pour déterminer les valeurs relatives, nous avons mis en relation les totaux figurant au bas du Tableau 4 (ligne «°Total prop.°») avec le nombre total des reformulations pour P1, P2 et P3. Les pourcentages de «°plus-value°» (PV), «°neutre°» (N) et «°moins-value°» (MV) par proposition figurent dans le Tableau 5.

Tableau 5. Pourcentages de PV, N et MV par proposition du Reformulateur

	Plus-value	Neutre	Moins-value
P1	27,3	15,2	57,6
P2	4,8	23,8	71,4
P3	16,7	25,0	58,3

Si l'on agrège les cas de «°plus-value°» et ceux d'une proposition «°neutre°», on constate que P1 et P3 fournissent un pourcentage de reformulations judicieuses pratiquement identique (42,5 % vs 42,7 %), générant par complémentarité un nombre de reformulations inappropriées aussi élevé l'un que l'autre (57,6 % vs 58,3 %). P2 se distingue par un taux de plus-values (23,8 %) extrêmement bas et une fréquence des moins-values particulièrement élevée (71,5 %).

Discourse and Writing/Rédactologie
Volume 35, 2025
http://journals.sfu.ca/dwr

Pour déterminer, dans le cas de propositions multiples, quelle proposition a la probabilité la plus forte de constituer une plus-value ou d'être tout simplement judicieuse, il a fallu extraire du Tableau 4 tous les segments ne donnant lieu qu'à une seule proposition (n = 16) et recalculer les pourcentages de PV, N et MV pour les propositions présentées par le Reformulateur en 1re, 2e et 3e positions. Les valeurs relatives ainsi obtenues sont reprises dans le Tableau 6.

Tableau 6. Pourcentages de PV, N et MV du Reformulateur en cas de propositions multiples

	Plus-value	Neutre	Moins-value
P1	25,0	15,0	60,0
P2	5,0	25,0	70,0
P3	16,7	25,0	58,3

Du Tableau 6, il ressort que :

- la probabilité d'obtenir une plus-value est la plus forte pour la proposition présentée par le Reformulateur en 1re position°;
- la probabilité d'obtenir une réponse judicieuse est quasiment égale pour les propositions présentées en 1re (25 + 15 = 40 %) et en 3e position (16,7 + 25 = 41,7 %)°;
- la probabilité d'obtenir une proposition inappropriée est plus élevée que celle d'obtenir une proposition judicieuse, quelle que soit la position.

Le dernier volet de l'analyse des résultats concerne la nature des problèmes rédactionnels traités par le Reformulateur. Le Tableau 7 donne un aperçu complet du traitement de ces problèmes, classés ici par ordre décroissant du nombre de propositions, plus-values, neutres et moins-values cumulées.

Discourse and Writing/Rédactologie
Volume 35, 2025
http://journals.sfu.ca/dwr

Tableau 7. Nature des problèmes rédactionnels détectés par le Reformulateur

	Plus-value	Neutre	Moins-value	Total
lexie terne	7	6	18	31
tournure passive	0	1	8	9
complication syntaxique	1	3	4	8
majuscule/minuscule	1	0	3	4
sélection de la construction verbale	0	0	3	3
ambiguïté syntaxique	2	0	1	3
feuille de style, charte graphique, terminologie du mandant	0	0	2	2
virgule	2	0	0	2
pb de progression logique	0	0	2	2
pb d'adaptation à la typologie textuelle	0	0	1	1
coquille	1	0	0	1
orthographe d'usage	0	1	0	1
pb lexical	0	0	1	1
Totaux	14	11	43	68
Pourcentages	21	16	63	
Moyenne				5,2

Du Tableau 7, il ressort que :

- le Reformulateur a fait des propositions pour onze catégories différentes de problèmes rédactionnels, à savoir (dans l'ordre d'apparition dans la grille d'évaluation) : la typographie, la ponctuation, l'orthographe d'usage, la grammaire *stricto sensu* (uniquement la tournure passive), la syntaxe, le lexique, l'univocité, la progression logique de la phrase, l'expressivité de la phrase (uniquement « lexie terne »), l'adaptation à la typologie textuelle, l'adaptation au guide de style du mandant et à l'éthos de l'énonciateur°;
- le Reformulateur n'a rien détecté concernant les trois catégories suivantes : cohérence structurelle, cohésion (dynamique de l'information) et idiomaticité ; nous reviendrons sur ce constat au point 5 « Interprétation des résultats » ;
- les trois problèmes qu'il détecte le plus (31 lexies ternes, 9 tournures passives et 8 complications syntaxiques pour une moyenne de 5,2 détections par catégorie de problèmes) sont également ceux pour lesquels il génère le plus de moins-values°;
- concernant les lexies ternes (= 46 % des 68 propositions), le Reformulateur fait preuve d'hypersensibilité en la matière, le conduisant à un haut degré d'hypercorrection :

Discourse and Writing/Rédactologie
Volume 35, 2025
http://journals.sfu.ca/dwr

Exemple 1

E1 : Avec humour, mais aussi sérieux, Maurice Barthélemy raconte comment il a découvert son hypersensibilité et la manière dont il en a fait sa force.

- Reform. P1 : Avec humour, mais aussi sérieux, Maurice Barthélemy raconte comment il a découvert son hypersensibilité et la manière dont il **l'a transformée en une** force.
- Reform. P2 : Avec humour, mais aussi sérieux, Maurice Barthélemy raconte comment il a découvert son hypersensibilité et la manière dont il **s'en est servi comme d'une** force.
- Reform. P3 : Avec humour, mais aussi sérieux, Maurice Barthélemy raconte comment il a découvert son hypersensibilité et la manière dont il **s'est servi de cette qualité comme d'une** force.

Le segment de E1 *dont il en a fait sa force* ne souffrait pas tant de l'emploi d'un verbe vicaire que de la méconnaissance de la construction verbale adéquate dans ce contexte, à savoir *tirer sa force de*. Les propositions 1 à 3, résultant d'une détection systématique des verbes passe-partout, en échouant à identifier la construction verbale correcte, alourdissent inutilement la phrase. La proposition 1 n'était pas loin d'une plus-value ou d'un résultat neutre moyennant la suppression de l'article indéfini : *dont il l'a transformée en force* et non **dont il l'a transformée en une force.*

- la tendance à l'hypercorrection est encore plus marquée à l'égard de la tournure passive :

Exemple 2

E6 : La suite sera également publiée sous la collection « Enquêteurs Nocturnes ».

Reform. **: On publiera** également la suite sous la collection « Enquêteurs Nocturnes ».

Le Reformulateur tourne la phrase par l'actif avec un sujet indéfini sans intérêt informationnel, alors que la tournure passive permettait précisément d'éviter l'expression de l'agent.

Exemple 3

E15 : Enfoui dans les confins oubliés d'un monde terraformé, l'Anneau Souverain est découvert par Frodon, un jeune technicien de maintenance habitant la station Sarn Prime, en orbite autour de Valinor 7, la géante gazeuse d'une splendeur mythique en plein cœur de la galaxie Aewen.

Reform. : Enfoui dans les confins oubliés d'un monde terraformé, Frodon, un jeune technicien de maintenance habitant la station Sarn Prime, en orbite autour de Valinor-7, la géante gazeuse d'une splendeur mythique en plein cœur de la galaxie Aewen, **découvre** l'Anneau Souverain.

Dans ce cas-ci, la transformation de la phrase passive produit une anacoluthe (*enfoui – Frodon*), entraîne un éloignement excessif du verbe par rapport à son sujet (*Frodon […] découvre*) et rompt complètement la progression thématique.

Discourse and Writing/Rédactologie
Volume 35, 2025
http://journals.sfu.ca/dwr

5. Interprétation des résultats

Sur la base des résultats présentés dans la section 4, l'hypothèse selon laquelle le Reformulateur suggérerait des propositions correctes pour les segments agrammaticaux où le Correcteur se limite à une alerte (HYP 1a) est largement infirmée par l'absence de corrélation entre les deux modes de révision et le taux élevé de moins-values.

L'hypothèse selon laquelle le Reformulateur apporte une plus-value dans les cas où le Correcteur reste silencieux ou fait des propositions inappropriées à propos d'une anomalie grammaticale (HYP 1b) est infirmée par le taux de moins-values du Reformulateur en la matière. Il n'empêche qu'en une occasion le Reformulateur a corrigé une anomalie grammaticale (en réalité, une coquille) qui avait échappé au Correcteur et aux réviseurs humains :

Exemple 4

E7 : [Vous n'en pouvez plus d]'écumer les livres et les sites internet à la rechercher du remède miracle qui l'[= votre bébé] expédiera au pays des rêves ?

Reform : [Vous n'en pouvez plus d]'écumer les livres et les sites internet à la **recherche** du remède miracle qui l'expédiera au pays des rêves ? (E7|S2|P1)

La coquille, qui avait aussi bien échappé aux réviseurs humains qu'au Correcteur, a été automatiquement corrigée dans la proposition du Reformulateur.

En matière de grammaire du texte, sans infirmer totalement l'hypothèse selon laquelle le Reformulateur prendrait le relais du Correcteur en proposant des solutions pertinentes en cas de problèmes de cohérence référentielle et de dynamique de l'information relevés par les réviseurs humains (HYP2), on retiendra que le Reformulateur n'a rien détecté en matière de dynamique de l'information et qu'il s'est montré peu sensible aux problèmes de cohérence référentielle (1 cas de levée d'ambiguïté syntaxique, avec 2 propositions réussies sur 3) :

Exemple 5

E6 : Le profil des victimes est le même, ce ne sont que des hommes [...].

Reform. P1 : Le profil des victimes se répète : ce sont **uniquement** des hommes [...]. (plus-value : écarte le risque d'interprétation pragmatique minorative)

Reform. P2 : Le profil des victimes se répète : ce sont **seulement** des hommes [...]. (moins-value : n'écarte pas le risque d'interprétation pragmatique minorative)

Reform. P3 : Le profil des victimes se répète : ce sont **toujours** des hommes [...]. (plus-value, *idem* P1)

Discourse and Writing/Rédactologie
Volume 35, 2025
http://journals.sfu.ca/dwr

Concernant le traitement des défauts d'efficacité communicationnelle, l'HYP 3, selon laquelle le Reformulateur apporterait des améliorations en matière de fluidité, d'idiomaticité et d'adaptation au contexte discursif là où le Correcteur reste inactif, est totalement infirmée en matière d'idiomaticité (aucune détection), largement infirmée en ce qui concerne l'adaptation au contexte discursif (une seule détection, avec une proposition non pertinente) et partiellement confirmée pour la fluidité grâce aux détections relatives à la progression logique de la phrase et l'expressivité de la phrase (cette dernière catégorie comprenant les nombreux cas de lexie terne, évoqués plus haut).

En cours d'analyse, nous avons observé que le Reformulateur n'avait fait aucune proposition visant à pallier des faiblesses ressortissant aux trois catégories suivantes : cohérence structurelle, dynamique de l'information et idiomaticité – trois dimensions textuelles pour lesquelles on pouvait légitimement anticiper une performance supérieure de la part d'un outil fondé sur l'intelligence artificielle générative, en raison de sa capacité annoncée à appréhender les phénomènes discursifs complexes et à produire des reformulations contextuellement adaptées. Pour pouvoir apprécier ce constat à sa juste valeur, nous avons comparé ce silence aux détections faites en la matière par les réviseurs humains et le Correcteur. Du Tableau 8, qui présente les résultats de cette comparaison, il ressort au moins deux informations.

Discourse and Writing/Rédactologie
Volume 35, 2025
http://journals.sfu.ca/dwr

Tableau 8. Focus sur les détections humaines et du Correcteur dans les trois catégories d'erreurs non traitées par le Reformulateur

			HUM	COR =	COR -
Cohérence structurelle	Harmonisation	orthographique	E3	E3	
		rupture de point de vue, de mode d'exposition	E14		E14
		autres éléments du texte	E7		E7
Cohésion (dynamique de l'information)	Continuité thématique	sélection du procédé de coréférence	E4, E11		E4, E11
		défaut de résonance	E3		E3
	Connexité	sélection du connecteur	E1, E7		E1,E7
	Progression	excès de densité informationnelle	E15	E15	
	Étagement	manque de parallélisme syntaxique	E13		E13
		pb de séquentialité	E10		E10
Idiomaticité	Prédictibilité des cooccurrents	dissonance expressive	E7, E15		E7, E15
		autre pb manque preédictibilité	E1, E2		E1, E2
	Adaptation discursive	trace d'oralité (inadapté à l'écrit)	E15		E15
		lexie/construction trop soutenue/familière	E11		E11
		autre pb d'adaptation à la typologie textuelle	E3, E7, E11, E14	E3	E7, E11, E14
	Adaptation au lecteur cible	*autre pb d'adaptation au profil du lecteur*	E12		E12
	Adaptation au guide de style du mandant et à l'éthos de l'énonciateur	feuille de style, charte graphique, terminologie du mandant	E3		E3

Tout d'abord, il se confirme que, contrairement au Reformulateur, tant les réviseurs humains que le Correcteur ont détecté des problèmes relevant de ces trois catégories. Le silence du Reformulateur en la matière peut dès lors bien être interprété comme une limite intrinsèque de l'outil.

Ensuite, il est notable que tous les problèmes qui ont été détectés par les réviseurs humains dans chacune des trois catégories considérées ont également fait l'objet d'une alerte de la part du Correcteur, même si seules 3 détections du Correcteur (sur 23) donnent lieu à une correction judicieuse.

Autrement dit, si le Correcteur s'est montré très efficace pour détecter les problèmes de cohérence structurelle, de dynamique de l'information et d'idiomaticité, sa capacité à proposer des solutions adéquates demeure limitée. C'est précisément dans ce contexte que l'on aurait pu attendre du Reformulateur qu'il joue un rôle complémentaire, en apportant des reformulations susceptibles de pallier les insuffisances du Correcteur. Les données issues de notre corpus ne permettent pas de conclure à un tel succès.

Évitons toutefois de considérer ces conclusions comme définitives et généralisables à d'autres types d'écrits, dans la mesure où les résultats obtenus dans le cadre de cette étude reflètent et les réponses et l'état de développement du Reformulateur à un instant donné. Il s'agit donc d'un

Discourse and Writing/Rédactologie
Volume 35, 2025
http://journals.sfu.ca/dwr

instantané expérimental, nécessairement limité, dans un contexte technologique en constante évolution.

Par ailleurs, les résultats obtenus et discutés ont été orientés par nos hypothèses initiales, fondées sur la complémentarité annoncée entre le Correcteur et le Reformulateur. Si cette complémentarité n'a pas été confirmée, tant s'en faut, par l'analyse statistique, et si l'approche quantitative a mis en évidence l'importance du taux de propositions relevant de la surcorrection (reformulations correctes, mais de qualité équivalente à la version initiale, désignées comme «°neutres°» dans notre protocole d'évaluation) ou, plus préoccupant, de l'hypercorrection (propositions entraînant une dégradation du texte initial, dites «°moins-values°»), une analyse qualitative plus approfondie des suggestions pertinentes du Reformulateur laisse entrevoir un réel potentiel de l'outil, notamment en matière de fluidité de l'expression (allègement syntaxique) :

Exemple 6

E9 : À des milliers de kilomètres de tout ce qu'il connaît, loin de son père Semyon, qui lui a tout appris, et de ses terres familières, Maksim découvre une nouvelle vie, pleine de promesses, mais aussi d'incertitude.

Ref. : À des milliers de kilomètres de **tous ses repères**, loin de son père Semyon, qui lui a tout appris, et de ses terres familières, Maksim découvre une nouvelle vie, pleine de promesses, mais aussi d'incertitude.

Par la nominalisation réussie de la proposition relative substantive, le Reformulateur allège un tant soit peu la syntaxe accumulative de la phrase, en intervenant au niveau d'un segment et d'une manière (pourtant élémentaire) qui avaient échappé à la détection humaine assistée par le Correcteur.

En définitive, si les hypothèses initiales ont été globalement infirmées, les résultats obtenus ne disqualifient pas pour autant le Reformulateur. Loin de devoir être écarté, cet outil mérite une attention renouvelée, à condition de déplacer le regard : son efficacité ponctuelle, observée dans certaines reformulations pertinentes, gagnerait à être explorée dans le cadre d'une étude qualitative ciblée, affranchie de la comparaison systématique avec la détection humaine assistée par le Correcteur.

En tout état de cause, à ce stade de l'exploration de l'outil, il apparaît clairement que le potentiel du Reformulateur ne peut être pleinement exploité qu'à condition d'être encadré par une expertise humaine capable d'en évaluer les apports comme les limites. Car, en l'absence d'une maîtrise préalable de la langue et d'un regard critique sur les suggestions générées,

Discourse and Writing/Rédactologie
Volume 35, 2025
http://journals.sfu.ca/dwr

l'utilisateur non averti risque de sélectionner à l'aveuglette des reformulations inappropriées, compromettant ainsi la qualité du texte initial — voire l'altérant davantage au lieu de l'améliorer.

6. Recommandations pédagogiques

Au terme de cette étude, notre démarche s'inscrit plus que jamais dans ce que McKenna et Tshuma (2025) identifient comme le «°défi de la double expertise° », concept emprunté à Punya Mishra, lequel consiste à articuler une connaissance approfondie d'un domaine avec une maitrise critique de l'intelligence artificielle. La traduction de cette vision, séduisante sur le plan des principes, en actions concrètes sur le plan pédagogique, nous amène à proposer les recommandations suivantes.

Premièrement, pour rendre les étudiants aptes à apprécier les propositions d'une IA générative à leur juste valeur, il va falloir investir d'urgence dans le renforcement des compétences rédactionnelles dans la langue de travail, en priorité dans les institutions d'enseignement où la population étudiante est de plus en plus internationale et fait preuve d'une maitrise langagière de moins en moins satisfaisante. Les institutions accueillant un public plus homogène du point de vue linguistique ne sont toutefois pas dispensées d'un tel investissement puisque les jeunes générations sont de moins en moins exposées à l'écrit, en particulier dans le registre de langue soutenue. Il ne suffira pas d'introduire dans le cursus un module de réécriture avec l'outil d'IA générative, à l'instar de ce qui a pu se faire dans les écoles de traduction ces dernières années vis-à-vis de la post-édition, pour atteindre le niveau de compétence linguistique permettant de «°faire mieux que la machine°».

Deuxièmement, dans la pratique pédagogique même, il faudra en permanence attirer l'attention sur le risque de biais d'ancrage pointé très justement par Bramati (2025). L'expérience avec la post-édition montre à suffisance à quel point les étudiants ont du mal à se détacher de l'impression positive globale suscitée par des traductions automatiques désormais dépourvues de grossières fautes de langue. En ce sens, une reproduction *in vivo*, à plus petite échelle, de l'expérience menée ici, permettrait de montrer aux étudiants comment un expert de la langue peut interagir avec le Reformulateur (ou d'autres applications d'IA générative du même type) : non pas en le considérant comme une autorité, mais en confrontant ses propositions aux savoirs établis dans les ouvrages et les outils de référence, en identifiant ses biais ou ses lacunes, et en s'en servant comme d'un levier pour enrichir l'expertise humaine, plutôt que pour la remplacer.

Troisièmement, dans les écoles de traduction, les enseignants pourraient préconiser le recours aux outils d'aide à la rédaction ou à la révision avec les mêmes réserves qu'ils formulent à l'égard

Discourse and Writing/Rédactologie
Volume 35, 2025
http://journals.sfu.ca/dwr

du dictionnaire bilingue, dont il ne faut pas trop attendre non plus. En effet, le traducteur professionnel ne consulte le plus souvent pas ce type de dictionnaire pour trouver la traduction d'un mot ou d'une expression, mais plutôt pour *re*trouver la traduction qu'il lui faut (qu'il a sur le bout de la langue ou qui lui échappe momentanément) ou bien pour trouver une piste à confirmer dans un dictionnaire monolingue ou un corpus. Il en va de même pour un module comme le Reformulateur d'Antidote : il constitue une ressource supplémentaire, susceptible de stimuler la réflexion ou de suggérer des reformulations pertinentes, mais dont l'usage exige un regard critique et un effort de documentation. En effet, comme le Reformulateur propose des solutions, sans en expliciter les fondements, il ne saurait se substituer aux outils informatifs. Bref, son efficacité dépend étroitement de la capacité de l'utilisateur à évaluer la pertinence des suggestions, à la lumière de ses propres compétences linguistiques et des ressources explicatives disponibles.

7. Conclusion

L'analyse exploratoire du module de reformulation intégré à Antidote révèle un potentiel encore limité pour une utilisation autonome dans un cadre de formation à la révision professionnelle. Si certaines reformulations témoignent d'une capacité à améliorer ponctuellement la fluidité syntaxique ou à corriger des erreurs passées inaperçues, l'outil se distingue surtout par une forte propension à l'hypercorrection. L'absence de corrélation significative avec les détections du Correcteur, combinée à un taux élevé de propositions inappropriées, invite à une prudence méthodologique et didactique.

Ces résultats, bien que circonscrits à un corpus spécifique et à un moment donné du développement technologique de l'outil, soulignent la nécessité d'un encadrement critique de son usage. Le Reformulateur ne saurait se substituer aux outils informatifs ni à l'expertise humaine, mais peut, dans une perspective formative, constituer un levier de réflexion linguistique, à condition d'être utilisé dans un cadre pédagogique structuré, fondé sur la confrontation des suggestions générées avec les normes établies et les ressources de référence.

En définitive, cette étude plaide pour une intégration raisonnée et critique des outils d'IA générative dans la formation des professionnels du langage, en cohérence avec les exigences d'une double expertise – linguistique et technologique – désormais incontournable dans les métiers de la rédaction, de la traduction et de la révision.

Endnotes

Discourse and Writing/Rédactologie
Volume 35, 2025
http://journals.sfu.ca/dwr

1. Dans un souci de lisibilité et de fluidité rédactionnelle, nous avons choisi d'utiliser le masculin générique. Ce choix ne reflète aucune prise de position normative sur les questions de genre et ne remet pas en cause l'engagement de notre institution en faveur de l'inclusion et de l'égalité.

Références bibliographiques

Anckaert, Ph. (2023). La compétence technique : la malaimée des concours et examens de traductionº? *Équivalences*, *50*(1-2), 115-143.

Anctil, D. (2023). L'éducation supérieure à l'ère de l'IA générative. *Pédagogie collégiale*, *36*(3), 66-76.

Antidote (logiciel). (2025). In *Wikipédia*. https://fr.wikipedia.org/wiki/Antidote_(logiciel)

Bramati, A. (2025). *Grammatica contrastiva per i traduttori dal francese. Le trappole della lingua*. Cortina.

Charolles, M. (1981). L'ordre de la signification. *Pratiques*, *Numéro spécial Colloque de Cerisy*, 49-64.

Commission européenne. (2022). *Référentiel de compétences de l'EMT*. Commission européenne. https://commission.europa.eu/document/download/b482a2c0-42df-4291-8bf8-923922ddc6e1_fr?filename=emt_competence_fwk_2022_fr

Druide informatique Inc. (2010, octobre 22). Antidote reconnu d'intérêt pédagogique en France. *Nouvelles (Communiqués de presse)*. https://www.druide.com/fr/nouvelles/antidote-reconnu-dinteret-pedagogique-en-france

Druide informatique Inc. (2025). Reformulation et attribution à une intelligence artificielle [FAQ]. *Antidote.info*. Reformulation et attribution à une intelligence artificielle https://www.antidote.info/fr/assistance/depannage/contenu-linguistique/reformulation-et-attribution-a-une-intelligence-artificielle

Druide informatique Inc. (5 décembre 2023a). Antidote offre la reformulation par IA générative [Nouvelles]. *Communiqués de presse*. https://www.druide.com/fr/nouvelles/antidote-offre-la-reformulation-par-ia-generative

Druide informatique Inc. (5 décembre 2023b). La reformulation et l'intelligence artificielle d'Antidote [Nouvelles]. *Blogue Antidote*. https://www.antidote.info/fr/blogue/nouvelles/reformulation-et-intelligence-artificielle-antidote

Dupont, D., Fauvaux, T., et Ghenet, M. (1994). *La dynamique de l'information. Éléments de grammaire textuelle*. De Boeck Duculot.

Discourse and Writing/Rédactologie
Volume 35, 2025
http://journals.sfu.ca/dwr

Kreutzer, J., Briakou, E., Agrawal, S., et Fadaee, M. 2025. *Déjà Vu : Multilingual LLM Evaluation through the Lens of Machine Translation Evaluation* (arXiv:2504.11829). arXiv preprint.

Lemaire, N. (2022). *La révision unilingue comme levier d'amélioration de la qualité des textes expographiques multilingues. Méthode et analyse d'un corpus de panneaux et de cartels explicatifs relevant du domaine des beaux-arts* [Thèse de doctorat, Université libre de Bruxelles]. https://dipot.ulb.ac.be/dspace/handle/2013/352263

Marie, B., Fujita, A., et Rubino, R. (2021). Scientific credibility of machine translation research : A meta-evaluation of 769 papers. *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing*, 7297-7306.

McKenna, S., et Tshuma, N. (2025). L'IA peut être un danger pour les étudiants : 3 mesures que les universités doivent prendre. *The Conversation (France)*. https://theconversation.com/lia-peut-etre-un-danger-pour-les-etudiants-3-mesures-que-les-universites-doivent-prendre-256647

Frœliger, N. (2025). Dix ans de compétences EMT. Bilan provisoire et évolutions possibles. In *Former des traducteurs et des interprètes : Des prérequis au marché du travail* (Inalco / Éditions des archives contemporaines, p. 151-162). https://doi.org/10.17184/eac.8021

Rudolph, J., Tan, S., et Tan, S. (2023). ChatGPT : Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning and Teaching*, *6*(1), 342-363.

Thivierge Côté, L. (2025, avril 23). *Courriel : « Moteur de reformulation »* [Communication personnelle].