

**CHATGPT POUR LA RÉTROACTION CORRECTIVE ÉCRITE INTERACTIVE EN
APPRENTISSAGE DU FRANÇAIS LANGUE SECONDE**

TAEGAN HOLMES

Thèse soumise dans le cadre des exigences du programme de
Maîtrise ès arts en Études du bilinguisme

Institut des langues officielles et du bilinguisme

Faculté des arts

Université d'Ottawa

© Taegan Holmes, Ottawa, Canada, 2024

Table des matières

1	Introduction	1
2	Recension des écrits	10
2.1	L'importance de la RCÉ	10
2.2	La RCÉ des enseignants	15
2.3	Les outils numériques pour la RCÉ	17
2.3.1	Les correcticiels	17
2.3.2	Les systèmes d'évaluation automatique de la rédaction	21
2.3.3	Les systèmes de tutorat intelligents	23
2.4	L'IA en enseignement-apprentissage des langues	27
2.4.1	Les robots conversationnels	27
2.4.2	Les traducteurs automatiques	29
2.4.3	L'intelligence artificielle générative	31
2.5	Les affordances et les limites de l'IAG pour la RCÉ	32
2.5.1	La capacité de ChatGPT pour la correction des erreurs grammaticales	32
2.5.2	L'importance des requêtes	35
2.5.3	La traduction automatique avec ChatGPT	39
2.5.4	La rédaction assistée par l'IAG	41
2.5.5	Les limites de l'IAG	43
2.6	Les perceptions de l'IAG en milieu éducatif	46
2.6.1	Les perceptions des enseignants	46
2.6.2	Les perceptions des apprenants	47
2.7	Synthèse critique	48
2.8	Questions de recherche	51
3	Méthodologie	52

3.1	Cadre théorique.....	52
3.2	Intervention et collecte de données.....	54
3.3	Analyse des données.....	58
4	Résultats.....	61
4.1	Résultats du questionnaire préintervention.....	61
4.2	Résultats de l'analyse des fils de discussion d'interaction avec ChatGPT.....	64
4.2.1	Analyse quantitative des fils de discussion.....	65
4.2.2	Analyse des requêtes formulées par les apprenants.....	68
4.2.3	Analyse des réponses fournies par ChatGPT.....	71
4.3	Résultats du questionnaire post-intervention.....	86
5	Discussion.....	91
5.1	QR1.....	91
5.2	QR2.....	93
5.3	QR3.....	99
5.4	Recommandations.....	101
5.5	Les limites de l'étude.....	104
6	Conclusion.....	106
7	Références.....	108
8	Annexes.....	120

Liste de tableaux

Tableau 1 - Quelques types de RCÉ (inspiré de Guénette, 2009, p. 110).....	11
Tableau 2 - Les forces et les limites du correcticiel Antidote (Ouellet, 2013, p. 98)	19
Tableau 3 - Tableau de synthèse des contextes de requêtes	35
Tableau 4 - Tableau de synthèse des données quantitatives des fils de discussion	66
Tableau 5 - Tableau de synthèse des nouvelles requêtes créées par les deux groupes de participants.....	68
Tableau 6 - Typologie pour classer les requêtes étudiantes	70
Tableau 7 - Typologie pour classer les réponses de ChatGPT.....	71
Tableau 8 - Tableau de synthèse des réponses évaluant l'affirmation « J'ai aimé utiliser ChatGPT pour corriger mon texte ».....	87

Liste de figures

Figure 1 - Exemple de conversation avec ChatGPT	3
Figure 2 - Exemple d'un parallèle linguistique relevé par ChatGPT	8
Figure 3 - Exemple d'une traduction par ChatGPT.....	9
Figure 4 - Le correcticiel Antidote	18
Figure 5 - Le message d'erreur d'Antidote signalant une rupture de phrase.....	20
Figure 6 - Le SÉA Criterion (Wang, 2015, p. 93).....	22
Figure 7 - Le STI eTutor (Heift, 2010, p. 444).....	24
Figure 8 - Le traducteur automatique Google Traduction	29
Figure 9 - Les réponses aux questions portant sur l'utilisation de l'IA pour la rédaction et la révision textuelle en français.....	63
Figure 10 - Une réponse dans la catégorie « Réponse attendue » à la première requête imposée.....	72
Figure 11 - Une réponse dans la catégorie « Réponse attendue » à la deuxième requête imposée.....	73
Figure 12 - Une réponse dans la catégorie « Clarification » liée à une requête de clarification	74
Figure 13 - Une réponse dans la catégorie « Clarification » liée à une requête de clarification	74
Figure 14 - Une réponse dans la catégorie « Clarification » liée à une requête de clarification	75
Figure 15 - Une réponse dans la catégorie « Socialisation avec le participant ».....	75
Figure 16 - Une réponse dans la catégorie « Liste » qui répond à une requête sollicitant une information additionnelle.....	76
Figure 17* - Une réponse dans la catégorie « Modification » liée à une requête de clarification	77
Figure 18 - Une réponse dans la catégorie « Autocorrection » liée à une requête de clarification	78
Figure 19 - Une réponse dans la catégorie « Concession » liée à une requête de clarification	78

Figure 20 - Une réponse dans la catégorie « Réponse autre » qui répond à une question autre (la définition d'un terme anglais)	79
Figure 21* - Une réponse dans la catégorie « Révision » liée à une requête de clarification	79
Figure 22 - Diagramme de l'alignement requête-réponse.....	80
Figure 23 - Extrait d'une réponse incorrecte où ChatGPT ne décrit pas correctement une modification apportée.....	82
Figure 24 - Extrait d'une réponse incorrecte où ChatGPT prétend avoir effectué une modification qui n'a pas été faite.....	82
Figure 25 - Une réponse incorrecte où ChatGPT repartage le texte avec des erreurs précédemment corrigées.....	83
Figure 26 - Une réponse incorrecte où ChatGPT commet une erreur de conjugaison	84
Figure 27 - Une réponse inappropriée où ChatGPT repartage une liste au lieu d'effectuer une modification.....	85
Figure 28 - Une réponse inappropriée où ChatGPT repartage le texte révisé sans inclure d'explications.....	86
Figure 29 - Diagramme circulaire des réponses à l'affirmation « Formuler des requêtes pour obtenir de la RCÉ a été facile »	89

Résumé

ChatGPT est un outil d'intelligence artificielle générative qui est en mesure d'effectuer la correction des erreurs grammaticales dans plusieurs langues (Coyne et collab., 2023 ; Fang et collab., 2023 ; Wu et collab., 2023) et de fournir des rétroactions personnalisées au sein d'échanges interactifs synchrones (Su et collab., 2023). Les étudiants ayant utilisé ChatGPT semblent avoir des sentiments majoritairement positifs envers l'outil (Shoufan, 2023). Partant du postulat que les capacités interactives de ChatGPT pourraient s'avérer utiles pour la rétroaction corrective écrite (RCÉ), cette thèse propose une étude empirique expérimentale qui porte sur l'utilisation de ChatGPT pour la RCÉ interactive en apprentissage du français langue seconde. Les participants à l'étude (n=22) ont répondu à un questionnaire portant sur leurs pratiques autocorrectives de révision textuelle. Ils ont ensuite soumis un texte à ChatGPT et ont interagi avec la machine pour solliciter une RCÉ interactive lors d'une intervention ponctuelle. Les fils de discussion d'interactions apprenant-machine ont été analysés au moyen d'une analyse taxonomique pour produire une analyse qualitative du contenu qui a donné lieu à l'élaboration d'une typologie des interactions (requêtes-réponses). Les participants ont aussi répondu à un questionnaire post-tâche portant sur cette expérience. Les résultats montrent que les pratiques de révision textuelle avec ChatGPT sont idiosyncrasiques ; différents types de requêtes ont été créés, mais surtout des requêtes sollicitant des précisions et des clarifications. Par ailleurs, on constate aussi différents degrés d'engagement parmi les participants. Toutefois, la majorité des participants ont déclaré avoir des perceptions positives de l'expérience. Les réponses de ChatGPT se sont majoritairement avérées correctes et appropriées, mais leur qualité a pu souffrir en raison d'une requête imprécise ou mal formulée par un apprenant. Enfin, on propose des recommandations pour optimiser l'utilisation de ChatGPT pour la RCÉ interactive.

Remerciements

J'aimerais avant tout remercier ma superviseuse de thèse, Professeure Marie-Josée Hamel, qui m'a épaulée tout au long de la réalisation de ce projet. Elle m'a toujours tenue informée des nouvelles opportunités me permettant d'améliorer mon profil de chercheuse et elle m'a poussée à les saisir. Ses commentaires et ses suggestions ont été une aide précieuse et m'ont permis de rédiger une thèse dont je suis fière.

Je suis très reconnaissante de l'aide inestimable des membres de mon comité de thèse ; Professeur Jérémie Séror et Professeure Valia Spiliotopoulos. Leurs idées et leur soutien ont été indispensables. J'aimerais aussi remercier Professeure Martine Rhéaume, la professeure qui m'a très gentiment permis d'intervenir dans son cours pour mener ma collecte de données.

Enfin, j'aimerais remercier ma famille ; mes parents et mon frère Calder, ainsi que tous mes chers amis, mais en particulier Keegan Camp et Helen Mauch, qui m'aident toujours à garder la tête froide et Oriana Sarah Robinson pour ses sentiments de solidarité.

Mes remerciements les plus sincères.

Taegan Holmes

1 Introduction

On se questionne depuis plus d'un demi-siècle sur la capacité des machines de « penser ». En 1950, Alan Turing a publié un article révolutionnaire portant sur ce sujet (Rossoni & ChatGPT, 2022) qui a marqué un moment clé dans l'histoire de l'informatique. Ensuite, pendant les années 1950, John McCarthy a créé le terme « intelligence artificielle » ou « IA » (Kushmar, Vornachev, Korobova & Kaida, 2022). Bien que l'IA ne soit pas une nouveauté, de récents progrès dans ce domaine ont fait que ses implications actuelles constituent un sujet en plein essor. Actuellement, l'intégration des outils d'IA, notamment les « grands modèles de langage » ou « GML » (Large Language Models ou LLM en anglais), en éducation est devenue un sujet qui soulève l'attention, surtout en enseignement des langues. Cette tendance est surtout en raison des « [...] avancées dans l'IA pendant les années récentes qui ont rendu plus accessibles les outils et les services d'aide à la rédaction qui ont radicalement changé le processus de la rédaction et de l'édition » (traduction libre de Godwin-Jones, 2022, p. 5). D'autre part, selon Moulieswaran et Prasantha Kumar, il apparaît que « l'IA comporte des affordances pour l'enseignement des langues » [traduction libre] (2023, p. 292).

Les GML font partie de l'IA générative (IAG) et sont des « [...] modèle[s] de langage constitué[s] par apprentissage profond à partir de mégadonnées » (Office québécois de la langue française, 2023). L'Office québécois de la langue française précise que les GML « [...] permettent à des agents conversationnels de répondre à des requêtes en langage naturel de façon syntaxiquement et sémantiquement correcte, en prenant en compte un grand nombre de paramètres, pouvant atteindre l'ordre du billion (10^{12}) » (2023, paragr. 1). Ces outils d'IAG deviennent plus accessibles et il est donc impératif de regarder d'un œil critique leurs impacts potentiels sur le milieu éducatif.

La présente étude se concentre sur ChatGPT, un GML développé par l'entreprise OpenAI. ChatGPT est un « transformateur génératif préentraîné » (generative pre-trained transformer ou GPT en anglais) capable d'interagir avec un utilisateur d'une manière conversationnelle (OpenAI, 2024b). « ChatGPT est un modèle apparenté à InstructGPT, qui est entraîné à suivre une instruction dans une requête et à répondre avec une réponse détaillée » (traduction libre de OpenAI, 2024b, paragr. 2). ChatGPT suit le même modèle conversationnel qu'InstructGPT, mais ce qui le rend différent de cet autre modèle est sa nature conversationnelle (OpenAI, 2024b). Grâce à cette nature interactive, l'utilisateur peut entrer en conversation avec ChatGPT.

On s'intéresse à ChatGPT en particulier parce que c'est un système qui attire beaucoup d'attention depuis son lancement en 2022 (Wu et collab., 2023). Aujourd'hui, il existe de nombreux modèles semblables à ChatGPT (Claude, Perplexity, Gemini, etc.), mais lorsque cette étude a été originalement conçue, ChatGPT était le point de mire en ce qui concerne les GML conversationnels.

Pour communiquer avec ChatGPT, un utilisateur envoie une requête contenant une question ou des directives et ChatGPT parcourt sa base de données et utilise des techniques de traitement du langage naturel (TLN) où l'outil transforme les données textuelles en données structurées (Godwin-Jones, 2022, p. 5) afin de générer une réponse. La conversation continue selon ce modèle (Figure 1). En tant que générateur de contenu, ChatGPT vise à produire la réponse la plus probable en utilisant un algorithme de prédiction textuelle (OpenAI, 2024c). La réponse la plus probable n'est pas toujours la réponse la plus correcte. Godwin-Jones affirme que

[l]es GML sont des systèmes d'IA construits à partir d'énormes collections de données analysées par l'apprentissage automatique et permettant de traiter le langage humain de manière de plus en plus efficace. Les modèles linguistiques utilisés dans le cadre du TLN ne sont pas

basés sur une connaissance grammaticale de la langue, mais plutôt sur une modélisation artificielle et mathématique des données linguistiques. [Traduction libre] (2022, p. 11)

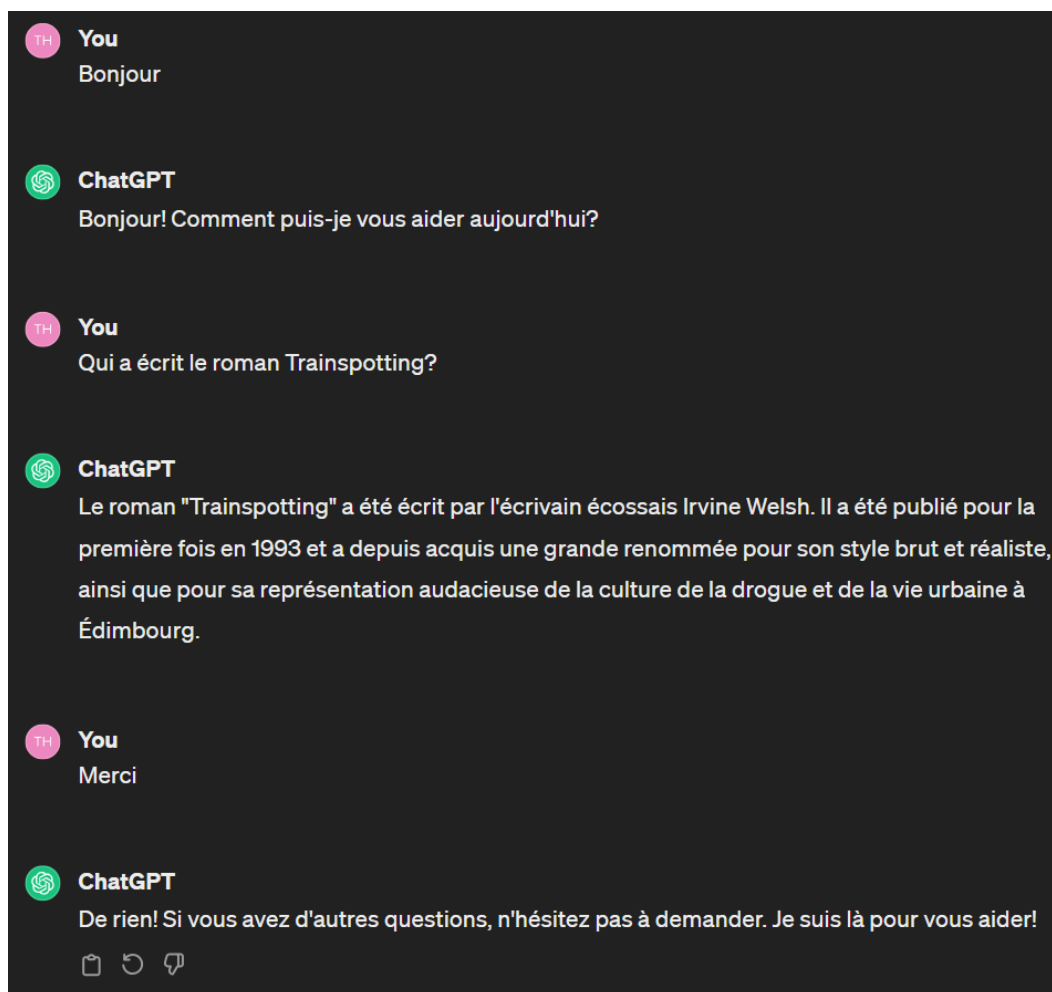


Figure 1 - Exemple de conversation avec ChatGPT

Bien que ChatGPT emploie un algorithme de prédiction textuelle pour générer ses réponses, l'outil a été entraîné avec une technologie qui s'appelle « l'apprentissage profond » et l'on a fait des ajustements avec le renforcement à partir de la rétroaction humaine (Reinforcement Learning From Human Feedback en anglais). En informatique, l'apprentissage profond est « [...] un sous-ensemble de l'apprentissage automatique, qui consiste essentiellement en un réseau neuronal de trois couches ou plus » (IBM, 2024, paragr. 1). ChatGPT

[a] été entraîné sur un grand ensemble de données pour créer un réseau neuronal d'apprentissage profond, lui permettant de comprendre les données textuelles et de générer des réponses semblables à celles des êtres humains en prédisant les mots suivants au sein d'une chaîne de mots donnée. (Traduction libre de Su, Lin & Lai, 2023, p. 2)

Dans le cadre de l'apprentissage profond, les réseaux neuronaux « [...] tentent d'imiter le cerveau humain à l'aide d'une combinaison d'entrées de données, de pondérations et de biais. Ces éléments fonctionnent ensemble pour reconnaître, classer et décrire avec précision les objets présents dans les données » (IBM, 2024, paragr. 10). Le renforcement à partir de la rétroaction humaine est une stratégie où « [...] un modèle d'IA apprend à modifier son comportement et à améliorer sa performance en se basant sur un modèle de récompense créé par la rétroaction humaine » (traduction libre de Su et collab., 2023, p. 2). Ainsi, grâce à ces différentes méthodes d'entraînement, ChatGPT va au-delà de la prédiction textuelle en formulant ses réponses.

OpenAI décrit que « leurs modèles textuels sont des outils avancés de traitement des langues qui peuvent générer, classifier et résumer le texte avec de hauts degrés de cohérence et de précision » (traduction libre de OpenAI, 2024c, paragr. 4). ChatGPT peut aussi corriger les textes et l'étude préliminaire de Wu, Wang, Wan, Jiao & Lyu (2023) suggère que ce dernier est un outil prometteur pour la correction des erreurs grammaticales. Pour cette thèse de maîtrise, on s'intéresse à la manière dont ChatGPT pourrait être intégré dans les pratiques de révision textuelle en apprentissage du français langue seconde (FLS).

Les études récentes tentent de démontrer que les transformateurs génératifs préentraînés comportent plusieurs affordances pour l'enseignement et l'apprentissage des langues, surtout en ce qui a trait à la traduction (Jiao et collab., 2023), la rédaction et la révision textuelle (Su et collab.,

2023) ainsi que la correction des erreurs grammaticales (Coyne, Sakaguchi, Galvan-Sosa, Zock & Inui, 2023 ; Fang et collab., 2023 ; Loem, Kaneko, Takase & Okazaki, 2023 ; Wu et collab., 2023).

Selon la théorie des affordances de Gibson, une affordance fait référence à la « [...] faculté de l'homme, et de l'animal en général, à guider ses comportements en percevant ce que l'environnement lui offre en termes de potentialités d'actions » (Luyat & Regia-Corte, 2009, p. 298). D'après les résultats des études empiriques, les principales affordances de ChatGPT en tant qu'outil d'aide à la rédaction semblent être une capacité élevée pour le rappel pendant les tâches de correction des erreurs grammaticales¹ (Coyne et collab., 2023 ; Fang et collab., 2023 ; Su et collab., 2023 ; Wu et collab., 2023), une performance comparable aux traducteurs automatiques de pointe, surtout en utilisant GPT-4 (la version payante de ChatGPT) (Jiao et collab., 2023) et sa nature adaptative (Su et collab., 2023). Ainsi, ces affordances nous apparaissent comme une occasion propice à intégrer ChatGPT dans un cours de FLS pour la révision textuelle, plus précisément pour la rétroaction corrective écrite (RCÉ).

Notre posture dans ce projet de recherche est motivée par le fait d'être à la fois une apprenante et une enseignante de FLS qui s'intéresse aux outils numériques pour la révision textuelle. Conséquemment, les affordances de ChatGPT pour la RCÉ sont d'intérêt pour nous, tout autant que pour de (futurs) enseignants et apprenants de FLS.

En ce qui concerne la RCÉ, il existe un débat portant sur son efficacité et sur son importance (Hamel & Bibeau, 2021), mais de nombreux chercheurs s'entendent sur son importance (Bitchener & Ferris, 2011 ; Godwin-Jones, 2022 ; Su et collab., 2023 ; Woodworth & Barkaoui, 2020). « La théorie et les recherches soulignent le rôle important de la RCÉ sur le développement en langue

¹ En correction des erreurs grammaticales, le rappel fait référence à la capacité d'un modèle de repérer correctement les erreurs grammaticales.

seconde et en particulier sur le développement de la production écrite en langue seconde » (traduction libre de Woodworth & Barkaoui, 2020, p. 235). Woodworth & Barkaoui avancent que « [...] le fait d'écrire fréquemment en langue seconde (L2) n'est pas suffisant ; les apprenants ont également besoin de recevoir des rétroactions afin de s'améliorer » [traduction libre] (2020, p. 235).

Les enseignants semblent être majoritairement d'accord sur l'importance de la RCÉ (Bitchener & Ferris, 2011 ; Hamel & Bibeau, 2021). Lors d'une étude menée en 2021 concernant les pratiques correctives des enseignants de langue, 90 % des enseignants sondés ont dit « [...] que la RCÉ est une tâche pédagogique très importante pour eux [...] » (Hamel & Bibeau, p. 6). Bitchener et Ferris (2011) constatent aussi que les enseignants de langue seconde croient à l'importance de fournir une RCÉ. En outre, Lee (2017) affirme que les théories cognitive, sociocognitive et socioculturelle s'entendent sur le rôle positif qu'occupe la RCÉ dans le développement de la production écrite en L2.

Les résultats de l'étude d'Hamel et Bibeau démontrent que « [...] la RCÉ est une tâche importante et fastidieuse pour les enseignants [...] » (2021, p.19). Fournir une RCÉ peut être un processus exigeant (Op. cit). Il existe des outils qui permettent d'appuyer les apprenants durant leur processus d'écriture et à recevoir une RCÉ, tels que les correcticiels (*Antidote* et *BonPatron*) et les systèmes d'évaluation automatique de la rédaction (SÉA) (tels que *Criterion* et *Pigai*²), mais ces outils n'ont pas été construits pour entrer en conversation avec les utilisateurs. La communication est en quelque sorte unidirectionnelle ; une fois le texte soumis, le correcticiel ou le SÉA l'analyse et offre des pistes de correction sous la forme de boîtes de dialogue qui invitent l'utilisateur à réfléchir

² Pour l'anglais seulement

sur ce qu'il a écrit et à réviser son texte. Ces outils ne sont cependant pas prévus pour permettre à l'utilisateur de demander des précisions ou de négocier le sens. Il existe aussi des systèmes de tutorat intelligents qui sont des outils numériques plus adaptatifs, mais qui peuvent être trop contraignants et sont limités dans la communication (Slavuj, Kovačić & Jugo, 2015).

En revanche, en raison des capacités interactives de ChatGPT, nous postulons que cet outil numérique pourrait s'avérer utile pour fournir en temps réel une RCÉ aux apprenants au sein d'échanges interactifs synchrones. « La fonction interactive de ChatGPT permet une collaboration humaine avec le GML pour améliorer sa performance [...] » (traduction libre de Bang et collab., 2023, p. 1). Su et collab. avancent « [qu']en tant qu'IAG, ChatGPT peut fournir des réponses et des rétroactions personnalisées » [traduction libre] (2023, p. 9). En outre, les études tentent de démontrer que ChatGPT est un traducteur automatique puissant (Jiao et collab., 2023), surtout pour les langues à ressources élevées et moyennes (Bang et collab., 2023). Il semble que ChatGPT est capable de traduire ses explications et de relever des comparaisons linguistiques lorsque demandé par un utilisateur, ce qui pourrait s'avérer utile pour les apprenants d'une langue seconde. Les Figures 2 et 3 sont des exemples créés qui illustrent ceci. Cependant, il semble que ChatGPT éprouve plus de difficultés avec les langues à faibles ressources (Bang et collab., 2023 ; Jiao et collab., 2023).

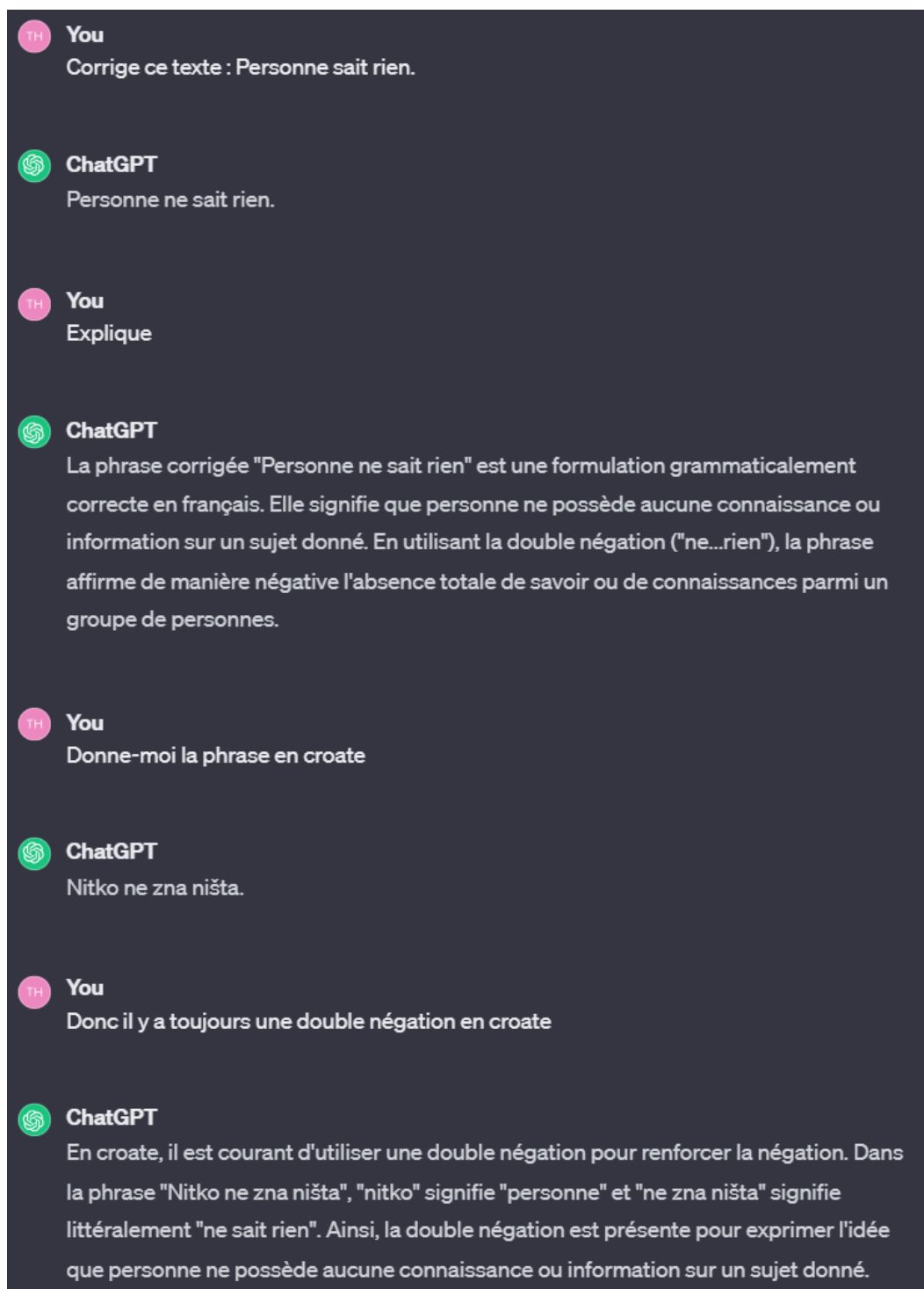


Figure 2 - Exemple d'un parallèle linguistique relevé par ChatGPT

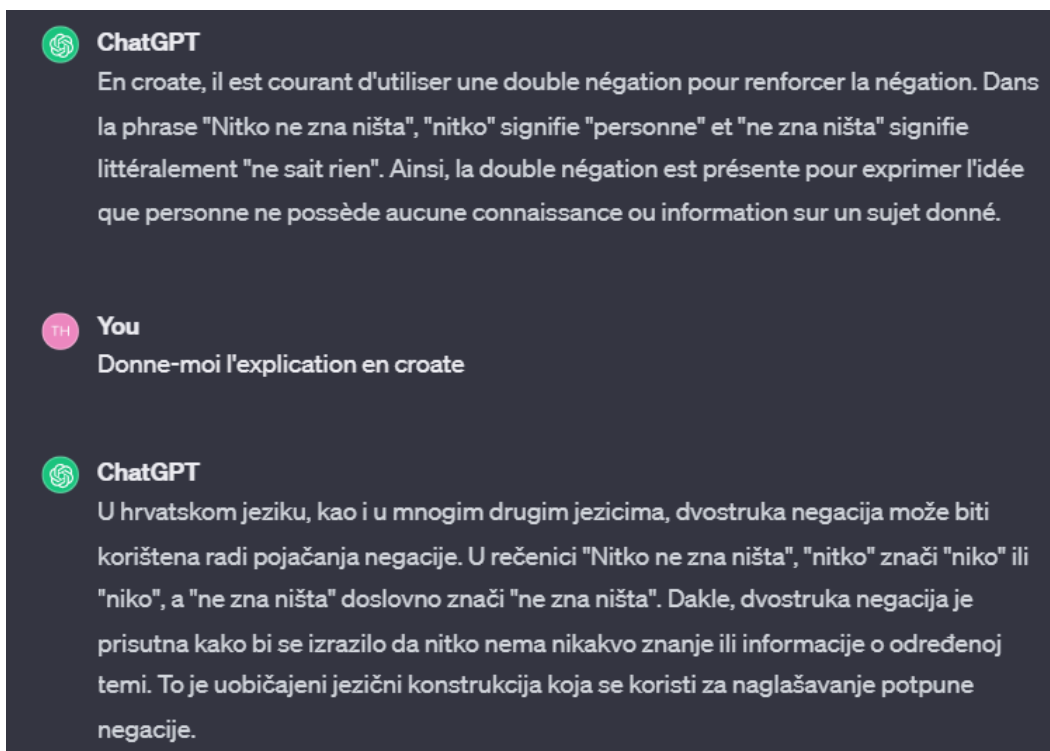


Figure 3 - Exemple d'une traduction par ChatGPT

Il existe à ce jour très peu de recherches menées auprès de l'IA générative pour la RCÉ et encore moins en français. À notre connaissance, il n'y a pas d'articles à ce jour publiés sur l'IAG pour la RCÉ en français. Nous souhaitons combler en partie cette lacune dans la littérature avec ces recherches.

2 Recension des écrits

Cette recension des écrits vise à dresser l'état des lieux des recherches pertinentes concernant l'utilisation de l'intelligence artificielle à des fins pédagogiques, ainsi qu'à justifier la présente étude. Les publications sur l'IAG représentent un domaine qui est actuellement très prolifique et conséquemment cette recension des écrits porte sur les articles publiés jusqu'en décembre 2023.

Dans ce chapitre, on s'intéresse d'abord à la RCÉ, à son importance et à la RCÉ des enseignants. Puis, on passe en revue des outils numériques qui permettent la RCÉ. On discute ensuite des études empiriques menées qui mettent en lumière les affordances de l'IAG pour fournir une RCÉ interactive, ainsi que ses limites. Ensuite, on mentionne les perceptions des enseignants et des étudiants envers l'IAG. En conclusion de ce chapitre, on propose une synthèse critique afin de donner une vue d'ensemble des recherches recensées et des applications qu'elles suggèrent. On présente aussi les questions de recherche qui motivent la présente étude.

2.1 L'importance de la RCÉ

La littérature sur l'importance de la RCÉ, en particulier pour les apprenants des langues secondes, est très vaste et l'on n'y retrouve pas de consensus définitif. En 1996, John Truscott a publié un article en réaction aux travaux de D. Ferris ayant pour but de remettre en question la valeur de la RCÉ, ce qui a été l'élément déclencheur pour le fameux débat « Truscott-Ferris » (Cf. Hamel & Bibeau, 2021). Dans ce débat, Ferris tentait de souligner la valeur de la RCÉ tandis que Truscott tentait de faire le contraire. Bien qu'il n'existe toujours pas de consensus sur la valeur de la RCÉ, cette dernière demeure « [...] partie intégrante de l'enseignement de l'écriture en langue seconde dans le monde entier » (traduction libre de Reinders & Mohebbi, 2018, p. 1). En outre, les recherches ont démontré que la RCÉ peut faciliter le traitement et la rétention des formes correctes pour le développement de l'écrit chez les apprenants (Ellis, 2009). D'ailleurs, on constate que la

RCÉ peut mener les apprenants à être plus autonomes en écriture (Ferris, 2003). Ainsi, son importance va au-delà de la facilitation de la maîtrise linguistique. Pour qu'un apprenant puisse s'améliorer en production écrite en L2, il doit non seulement pratiquer et s'exercer en écrivant, mais aussi recevoir des rétroactions sur ses écrits (Barkaoui, 2007 ; Woodworth & Barkaoui, 2020), ce qui souligne l'importance de la RCÉ.

Type de rétroaction	Description
Directe	Une RCÉ qui corrige directement l'erreur.
Indirecte	Une RCÉ qui signale l'erreur sans la corriger.
Métalinguistique	Une RCÉ présentant une explication métalinguistique de l'erreur.
Sélective/ciblée	Une RCÉ ciblant quelques types d'erreurs dans un texte.
Compréhensive/ concentrée/exhaustive	Une RCÉ ciblant toutes les erreurs dans un texte.

Tableau 1 - Quelques types de RCÉ (inspiré de Guénette, 2009, p. 110)

Il existe de nombreuses manières de classifier les différents types de RCÉ (Guénette, 2009). Le Tableau 1 en illustre quelques-uns. Les chercheurs avancent différentes théories concernant le type de RCÉ le plus efficace (Lee, 2020). Certains chercheurs postulent que la RCÉ sélective (lorsqu'on ne signale/corrige que des erreurs ciblées) est plus efficace que la RCÉ compréhensive (où l'on signale/corrige toutes les erreurs), mais d'autres l'ont démontré autrement. Scarcella (1996) postule que la RCÉ sélective pourrait mener à la fossilisation des erreurs non traitées. Les résultats de l'étude de Şakrak-Ekin & Balçikanli (2019) démontrent que ce sentiment est aussi partagé par certains enseignants. De plus, McMartin-Miller (2014) souligne qu'il est difficile de choisir sur quels types d'erreurs il faut se concentrer pendant la RCÉ sélective, ainsi que le nombre d'erreurs de divers types qui devraient être traitées à la fois. « Un autre facteur à considérer lors de la prise

de décisions pédagogiques concernant le type de RCÉ à fournir est le rôle des différences individuelles entre les apprenants » (traduction libre de Yilmaz & Granena, 2015, p. 149).

Ferris (2010) postule que d'une perspective rédactionnelle, se concentrer sur un ou deux éléments semble être trop peu et que se concentrer sur quinze à vingt éléments est probablement trop. En revanche, d'autres chercheurs comme Lee (2020) constatent qu'il est plus efficace de fournir de la RCÉ sur un seul élément linguistique que de la fournir pour quatre à cinq. Ellis, Sheen, Murakami & Takashima avancent que cela est parce qu'un

[...] grand nombre de corrections portant sur un ensemble varié de phénomènes linguistiques (et peut-être aussi sur des questions de contenu et d'organisation) n'est pas susceptible de favoriser la prise de conscience et la cognition qui peuvent être nécessaires pour que la RCÉ fonctionne en matière d'acquisition. En revanche, une correction ciblée de manière répétée sur un problème grammatical très précis peut avoir un effet plus significatif, comme l'ont montré des études sur l'apprentissage des langues à l'oral. [Traduction libre] (2008, p. 368)

L'efficacité des différents types de RCÉ (directe, indirecte, métalinguistique) pour la RCÉ concentrée/compréhensive est non concluante (Buckingham & Aktuğ-Ekinci, 2017 ; Shintani, Ellis & Suzuki, 2014 ; Van Beuningen, De Jong, & Kuiken, 2012), mais on constate que la RCÉ fournie en mode synchrone semble être bénéfique (Arroyo & Yilmaz, 2018 ; Shintani, 2016 ; Shintani & Aubrey, 2016). En effet, la rétroaction corrective immédiate s'est avérée bénéfique concernant la précision grammaticale lorsque fournie par les enseignants ou par le biais du clavardage (Arroyo & Yilmaz, 2018). Pendant l'étude d'Arroyo & Yilmaz (2018), les étudiants ayant reçu des rétroactions synchrones ont mieux performé que leurs pairs ayant reçu des rétroactions différées dans un test de production orale. Cependant, ces chercheurs n'ont pas

constaté de différence significative entre les étudiants ayant reçu des rétroactions synchrones et ceux ayant reçu des rétroactions asynchrones pendant un test de jugement grammatical.

Les résultats d'une autre étude menée auprès des étudiants d'anglais langue seconde dans une université japonaise démontrent que les rétroactions synchrones ont mieux facilité l'autocorrection que les rétroactions asynchrones pendant une tâche de rédaction (Shintani, 2016). Shintani & Aubrey (2016) ont aussi découvert durant une autre étude comparant la RCÉ synchrone et asynchrone que la RCÉ synchrone s'est avérée plus efficace pour l'amélioration de la précision des apprenants d'anglais langue seconde sur l'emploi du conditionnel hypothétique pendant un post-test immédiat. Chang, Cunningham, Müge Satar & Strobl (2018) avancent que cette amélioration s'est produite en raison de l'immédiateté la RCÉ synchrone, ce qui a appuyé l'apprentissage des notions grammaticales et a mené à moins de structures fossilisées. Cependant, Shintani & Aubrey (2016) ont constaté que les effets de la RCÉ synchrone ont semblé diminuer au fil du temps et les apprenants ont obtenu des scores moins élevés pour le post-test différé. « Autrement dit, si la rétroaction corrective synchrone s'est avérée efficace pour améliorer l'emploi précis de la structure ciblée, cette efficacité a diminué avec le temps » (traduction libre de Shintani & Aubrey, 2016, p. 309).

Enfin, selon la théorie socioculturelle de Vygotsky (1987), pour que la RCÉ soit utile pour un apprenant, elle doit être adaptée selon sa zone proximale de développement (ZPD) (la distance entre ce qu'un apprenant peut accomplir seul et ce qu'il peut accomplir avec de l'aide).

La théorie socioculturelle

[...] part du principe que tout le développement cognitif, y compris le développement de la langue, est le résultat d'interactions sociales entre les individus, en particulier lorsque les

apprenants ont la possibilité de collaborer et d'interagir avec des locuteurs de la langue cible plus compétents qu'eux (par exemple, des enseignants et des apprenants plus avancés).

(Traduction libre de Bitchener & Ferris, 2011, p. 18)

Zhang (2017) souligne aussi le fait que pour tirer profit de la RCÉ, il faut que les étudiants soient engagés de manière efficace avec cette dernière.

Avant de discuter des différentes sources de la RCÉ, il importe de mentionner les erreurs commises fréquemment par les apprenants de FLS qui peuvent solliciter des rétroactions. Il faut être conscient des types d'erreurs fréquentes afin de comprendre les types de RCÉ qui peuvent être fournis pour y remédier ; surtout afin d'évaluer si un outil numérique est en mesure de détecter ces types d'erreurs, de les signaler correctement et d'offrir des pistes de suggestion pour les rectifier.

Durant une étude menée en 2003, Mougeon & Mougeon ont identifié plusieurs erreurs structurelles dans les écrits des apprenants du FLS. Parmi ces erreurs structurelles, on retrouve : les erreurs portant sur l'emploi des verbes, des prépositions, des déterminants, des temps et des modes verbaux, les pronoms personnels objets, ainsi que d'autres erreurs grammaticales (tel que l'emploi du mauvais genre grammatical), l'omission de l'adverbe négatif « pas » et des erreurs sociolinguistiques et pragmatiques (tel que l'emploi du pronom « tu » au lieu du pronom « vous ») (Cf. Louis-Charles, 2012).

Hamel & Milčević (2007) ont analysé un corpus de textes rédigés par des apprenants de FLS dans un cours universitaire de 3e année afin d'analyser la fréquence de différents types d'erreurs lexicales. Le corpus compilé comportait 183 textes pour un total d'environ 50 000 mots. Les auteurs ont classé les erreurs lexicales dans le corpus selon trois classes ; erreurs de sens, erreurs de forme et erreurs de cooccurrence, chacune comportant plusieurs sous-classes. Les résultats de

cette étude ont relevé que « [l]es erreurs lexicales de la classe sens sont majoritairement dominante dans le corpus d'apprenants, avec un taux de 63 %. » (Hamel & Milčević, 2007 p. 38). Les erreurs de la classe cooccurrence avaient un taux de 25 % et les erreurs de forme avait un taux de 12 %. Parmi les erreurs de la classe sens, 21 % étaient des emplois des lexies de sens proche au lieu de la lexie visée (Hamel & Milčević, 2007). Un exemple de ce type d'erreur mentionné dans l'article est l'emploi de la lexie « nautiques » lorsque le contexte exige la lexie « aquatiques », tel que dans le segment « faites des exercices nautiques* » (Hamel & Milčević, 2007 p. 32).

Les erreurs susmentionnées ne constituent pas une liste exhaustive de toutes les erreurs commises par les apprenants de FLS, mais démontrent plutôt les erreurs fréquentes recensées et qui peuvent nécessiter une révision ou une RCÉ. Il est important pour les outils numériques (tel que ChatGPT) d'être en mesure de repérer ces types d'erreurs dans les textes. Ainsi, il faut mener des études empiriques à cet effet.

2.2 La RCÉ des enseignants

Tel que dit plus haut, il n'existe pas de consensus sur l'efficacité des divers types de RCÉ, cependant, la recherche souligne qu'elle est nécessaire pour le développement de la production écrite chez les apprenants (Woodworth & Barkaoui, 2020). La RCÉ provient souvent des enseignants. Les enseignants de langue sont majoritairement d'accord sur l'importance de la RCÉ (Bitchener & Ferris, 2011 ; Hamel & Bibeau, 2021) et « [...] ils ont presque tous, sous une forme ou une autre, des pratiques de RCE » (Hamel & Bibeau, 2021, p. 3). Tel que résumé dans le Tableau 1, les annotations ou les indices de la RCÉ peuvent prendre la forme de

[...] corrections directes, d'explications (métalinguistiques, commentaires) et/ ou d'indices. Les indices pourront être implicites (marquage du texte : soulignement, encerclement, etc.) ou explicites (usage de codes correspondant à des types d'erreurs). Ces divers types d'indications

ont pour but d'inciter l'apprenant à réaliser et comprendre ses erreurs pour l'amener à améliorer la qualité de sa production écrite. (Hamel & Bibeau, 2021, p. 2)

Bien que les enseignants de langue reconnaissent l'importance de la RCÉ (Op. cit), « [...] les recherches démontrent que la RCÉ des enseignants peut être inconsistante, non systématique, peu claire, imprécise, trop critique et/ou différée » (traduction libre de Woodworth & Barkaoui, 2020, p. 242). Lim et Phua (2019) affirment que superviser le processus de la rédaction en offrant des rétroactions utiles aux étudiants exige trop de temps, beaucoup de travail et s'agit d'un processus subjectif. Plusieurs chercheurs ont aussi noté que la RCÉ des enseignants est souvent différée (Arroyo & Yilmaz, 2018 ; Nazari, Shabbir & Setiawan, 2020 ; Woodworth & Barkaoui, 2020) parce que la correction des erreurs est une tâche qui prend du temps (Lee, 2003). Conséquemment, les enseignants ne sont souvent pas en mesure de fournir de la RCÉ synchrone bien que cette dernière s'avère souvent plus utile que la RCÉ asynchrone ou différée (Arroyo & Yilmaz, 2018 ; Shintani, 2016 ; Shintani & Aubrey, 2016). De nombreuses études ont trouvé que les apprenants d'une L2 préfèrent et demandent de la RCÉ fréquente et immédiate sur de nombreux brouillons (Woodworth & Barkaoui, 2020), ce qui représente un défi pour un enseignant parce que fournir une RCÉ de qualité exige du temps et des efforts non négligeables (Op. cit).

De plus, l'étude d'Arroyo & Yilmaz (2018) a révélé que « [...] la rétroaction différée n'est pas aussi efficace que la rétroaction immédiate pour une tâche où les apprenants doivent être précis en se concentrant sur le sens (tel qu'un test de production orale) » [traduction libre] (p. 967). L'étude de Zhang & Hyland (2018) menée pour comparer la RCÉ des enseignants et celle des SÉA a démontré qu'une apprenante en particulier avait préféré utiliser un SÉA parce qu'elle a pu obtenir des rétroactions immédiatement et elle a pu resoumettre son travail treize fois, ce qui a promu son autonomie.

2.3 Les outils numériques pour la RCÉ

Différents outils numériques ont été développés permettant d'appuyer les apprenants et les enseignants dans le processus de révision textuelle et ainsi servir à alléger la charge de travail de ces derniers. Parmi ces outils, on retrouve entre autres les correcticiels, les systèmes d'évaluation automatique de la rédaction (SÉA) et les systèmes de tutorat intelligents (STI). Ces outils ne sont pas nouveaux et existent depuis un bon moment, toutefois, dans le cadre de cette thèse, on s'intéresse aux versions modernes de ces outils. Cette section de la recension des écrits vise à mettre d'avant les affordances et les limites de ces trois types d'outils numériques.

2.3.1 Les correcticiels

Les correcticiels sont des logiciels qui signalent et proposent des corrections pour les erreurs d'orthographe et de grammaire (Figure 4). Leurs rétroactions sont quasi immédiates et un utilisateur peut resoumettre un même texte plusieurs fois. « Ces outils d'aide à la rédaction et à la correction ont été développés dans le but de prendre en charge non seulement la correction orthographique et grammaticale, mais aussi l'amélioration des textes sur le plan sémantique. » (Talbi & Legros, 2018, p. 7) Deux correcticiels populaires pour la langue française sont *BonPatron* et *Antidote*.

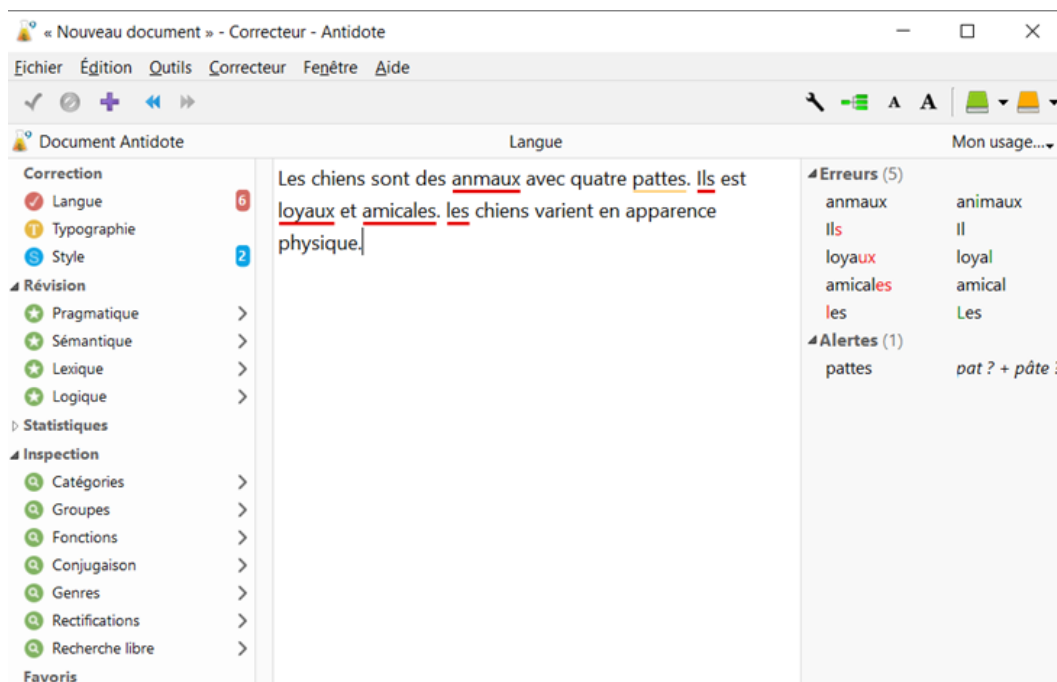


Figure 4 - Le correcticiel Antidote

« Antidote est un correcticiel (logiciel d'aide à la correction) qui combine de nombreux dictionnaires, des guides grammaticaux et des filtres intelligents pour travailler l'orthographe, la grammaire, la conjugaison et la syntaxe » (Hebbinckuys, Hong & Lory, 2021, p. 9). Hebbinckuys, Hong & Lory avancent que le correcticiel « Antidote représente une aide précieuse pour soutenir les étudiant·e·s dans la rédaction et la correction de textes écrits en français » (2021, p. 2). Cependant, ces chercheurs affirment qu'Antidote « [...] s'adresse avant tout à un public qui maîtrise déjà bien le métalangage employé dans les définitions ou les suggestions de corrections proposées [...] » (Hebbinckuys et collab., 2021, p. 2), ce qui signifie que les utilisateurs doivent être en mesure de comprendre la terminologie employée dans les explications des corrections proposées.

Forces	Limites
Repère les erreurs d'inattention	Ne comprend pas le texte qu'il corrige
Détecte une grande variété d'erreurs	Ne détecte pas toutes les erreurs
Met à notre disposition, facilement et rapidement, plusieurs ouvrages de référence	Peut suggérer une erreur
Suscite notre réflexion par le contenu des infobulles (corrections suggérées, notions de français, questions, etc.)	Demande certaines connaissances grammaticales pour comprendre le contenu des infobulles

Tableau 2 - Les forces et les limites du correcticiel Antidote (Ouellet, 2013, p. 98)

Ouellet (2013) résume les forces et les limites d'Antidote dans un tableau de synthèse (Tableau 2). Les recherches empiriques démontrent qu'Antidote peut détecter une grande variété d'erreurs (Hebbinckuys et collab., 2021 ; Ouellet, 2013 ; Talbi & Legros, 2018), toutefois, il ne détecte pas toujours toutes les erreurs présentes dans un texte. Par exemple, dans le projet de recherche de Mireault (2009) visant à comparer la détection des erreurs par Antidote et par les correcteurs humains, on constate que

[d]e façon générale, le correcticiel s'avère plus constant que les correcteurs humains en ce qui concerne l'orthographe d'usage et il détecte un total de 435 erreurs là où les correcteurs humains en signalent 322. Il s'avère toutefois moins efficace que les correcteurs humains pour les catégories « orthographe grammaticale », « vocabulaire » et « ponctuation ». (Ouellet, 2013, p. 75)

Les correcticiels peuvent aussi se tromper pendant la correction et peuvent désigner une forme correcte comme une forme erronée. Conséquemment, selon Hebbinckuys et collab. (2021), en utilisant Antidote, « [...] lorsque le logiciel suggère des corrections, l'utilisateur·trice doit porter un regard critique pour déterminer si celles-ci sont justifiées, nécessaires ou superflues » (p. 2).

Une autre limite d'Antidote identifiée par Ouellet est que le correcticiel ne comprend pas le texte qu'il corrige et « [...] tout ce qui n'est pas reconnu par le système est traité comme une erreur. Ainsi, Antidote indique tous les contextes où il se montre incapable de produire une analyse de la phrase » (2013, p. 59) comme des erreurs ou des ruptures de phrase (Figure 5).

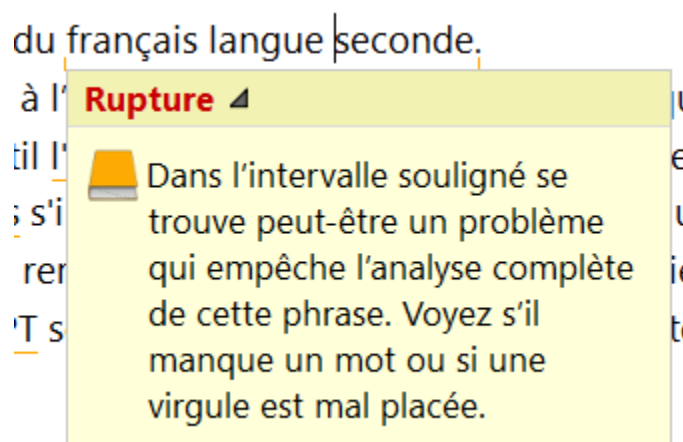


Figure 5 - Le message d'erreur d'Antidote signalant une rupture de phrase

BonPatron est un correcticiel qui « [...] a été conçu pour des usagers anglophones qui étudient le français ou le parlent comme langue seconde » (López Santiago, 2016, p. 51). Tout comme Antidote inclut plusieurs ouvrages de référence (Op. cit), BonPatron contient également des ressources additionnelles qui incluent des guides linguistiques, un outil de conjugaison et des dictionnaires (López Santiago, 2016).

Les explications des corrections de BonPatron utilisent moins de métalangage que celles d'Antidote et López Santiago explique qu'en passant la souris sur les erreurs détectées, les boîtes de dialogue contenant « [...] des explications faciles à comprendre et des exemples de bon usage » (2016, p. 53) sont affichées. Cependant, « [i] est important de souligner que ce logiciel ne reconnaît pas toujours les fautes commises par les apprenants, car il identifie les erreurs à l'aide

de règles préétablies, en fonction des erreurs repérées dans des textes soumis précédemment » (López Santiago, 2016, p. 54).

En bref, on constate que les correcticiels sont utiles pour attirer l'attention d'un utilisateur sur « [...] certaines formes potentiellement fautives, en formulant, dans certains cas, des propositions de correction. Le scripteur (ou l'apprenant) doit en faire l'évaluation avant de les accepter ou de les rejeter » (Ouellet, 2013, p. 61), ce qui pourrait constituer un défi pour les apprenants moins expérimentés ou formés à l'utilisation de ces outils (Hebbinckuys et collab. ; Ouellet, 2013). Comme on a dit plus haut, ces outils n'ont pas été construits pour entrer en interaction avec l'utilisateur, mais avec les récentes avancées en IAG, on commence à y intégrer des fonctionnalités génératives (Druide informatique, 2023 ; Grammarly, 2024).

2.3.2 Les systèmes d'évaluation automatique de la rédaction

Les systèmes d'évaluation automatique (SÉA) de la rédaction sont semblables aux correcticiels, mais ont été construits pour tenir compte du contexte discursif d'un texte écrit et peuvent donc offrir des pistes de suggestion pour l'orthographe et la grammaire, ainsi que pour l'organisation textuelle (Woodworth & Barkaoui, 2020) (Figure 6). Woodworth & Barkaoui (2020) soulignent qu'il ne faut pas confondre les correcticiels et les SÉA, toutefois, avec les avancées en IAG, on voit que les frontières entre ces deux types de systèmes s'estompent (Op. cit).

À l'instar des correcticiels, les SÉA proposent une RCÉ quasi immédiate et les textes peuvent être resoumis et recorrectés plusieurs fois.

The screenshot displays the ETS Criterion interface for a student named MisiaChen. The top navigation bar includes the ETS logo, the student's name and ID (1098161120), the submission date (October 26, 2010), and the score (Earn Money). Below this is a 'Trait Feedback Analysis Menu' with tabs for Grammar, Usage, Mechanics, Style, and Organization & Development. The 'Mechanics' tab is active, showing a list of error types on the left, with 'Missing Final Punctuation' selected. The main content area shows a sample letter with several errors highlighted in blue, corresponding to the selected category. A 'View Question' link is also visible.

Figure 6 - Le SÉA Criterion (Wang, 2015, p. 93)

L'usage réflexif de la RCÉ des SÉA (Yannakoudakis, Andersen, Geranpayeh, Briscoe & Nicholls, 2018) promeut l'autonomie des apprenants (El Ebyary & Windeatt, 2010). Ranalli, Link & Chukharev-Hudilainen (2017) postulent aussi que même si un SÉA se trompe durant la correction, les apprenants peuvent toujours profiter de la RCÉ incorrecte parce qu'elle attire leur attention sur les formes linguistiques, ce qui pourrait faciliter l'acquisition. En tant qu'outils numériques de correction linguistique, les SÉA peuvent d'autant plus alléger la charge de travail des enseignants (Tang & Rich, 2017) concernant la révision textuelle.

On constate que les SÉA comportent aussi des limites. Godwin-Jones avance qu'il n'existe « [...] aucun paramètre dans ces outils permettant de varier les rétroactions selon les caractéristiques des utilisateurs, tel que le fait d'être un apprenant de L2 » (2022, p. 9). Ranalli (2018) précise que la nature de la RCÉ automatique (RCÉA) se base sur une approche universelle qui ne prend pas en compte les différences entre les utilisateurs, tel que leurs niveaux de compétence en L2, leurs langues maternelles, leurs compétences en production écrite ou leurs formations scolaires. Les

informations fournies concernant les erreurs détectées sont plutôt déterminées par les capacités technologiques d'un SÉA que par des considérations pédagogiques (Ranalli, 2018). Ranalli décrit aussi la nature de la RCÉA comme étant générique et il précise qu'elle considère très peu (ou pas du tout) les différences individuelles (2018). Woodworth & Barkaoui (2020) affirment que les SÉA modernes ne peuvent pas adapter leurs rétroactions selon les besoins, les niveaux de compétences en L2 ou les buts des apprenants. Ranalli (2018) suggère que pour améliorer les SÉA, les développeurs devraient essayer de rendre plus précises les rétroactions apportées par le système (par exemple, en indiquant où une phrase mal construite pourrait être divisée avec un point-virgule ou un point final au lieu de fournir une rétroaction générale et vague indiquant simplement que la phrase est trop longue).

Une étude menée auprès des étudiants taiwanais a révélé que ces derniers avaient de mauvaises perceptions du SÉA *My Access!* parce que la RCÉA a été trop vague et générale pour leur permettre de se corriger et conséquemment, ils ont préféré obtenir des rétroactions de leurs pairs (Koltovskaia, 2020). De plus, les SÉA n'ont pas été construits pour échanger avec les utilisateurs et il n'est conséquemment pas possible de solliciter des précisions concernant les corrections proposées dans la RCÉA. Ainsi, comme avec les correcticiels, la communication reste essentiellement unidirectionnelle.

2.3.3 Les systèmes de tutorat intelligents

Le dernier outil numérique pour la RCÉ sur lequel on s'attarde est le système de tutorat intelligent (Ji, Han & Ko, 2023 ; Jia, Sun, Ma & Looi, 2022 ; Kushmar et collab., 2022 ; Ruan et collab., 2021). Bien que les systèmes de tutorat intelligents ou STI soient des outils issus de l'IA, on les mentionne dans cette sous-section parce qu'une des principales fonctionnalités de ces outils est leur capacité à proposer une RCÉ. Les STI sont des outils de technologie de l'information et de la

communication (TIC) qui fournissent aux apprenants de l’instruction ainsi que des rétroactions.

« Les STI sont des systèmes d’enseignement informatiques qui possèdent un contenu sous forme de base de connaissance (qui spécifie ce qui doit être enseigné), des stratégies d’enseignement (qui spécifient la manière d’enseigner ce contenu) ainsi qu’une connaissance sur le niveau de l’apprenant dans le contenu, afin d’adapter dynamiquement leur enseignement » (traduction libre de Zouaq, 2005, p. 5).

The screenshot shows the 'the eTutor' German Language Tutoring interface. At the top left, the logo 'the eTutor' is displayed with 'GERMAN LANGUAGE TUTORING' underneath. To the right is a German eagle emblem. Below the logo is a navigation menu with 'INTRODUCTION', 'CONTENTS', and 'DICTIONARY' tabs. The main content area is titled 'Kapitel 3 - Familie und Freunde'. It contains an introductory paragraph, a paragraph about a family holiday, and a listening exercise with a photo of a family picnic. The interface also has a top navigation bar with 'INTRODUCTION', 'CONTENTS', and 'DICTIONARY' tabs, and a 'Kapitel 3' indicator.

Figure 7 - Le STI eTutor (Heift, 2010, p. 444)

Les STI sont des modèles adaptables qui « [...] font des inférences sur les compétences d’un apprenant en lien avec un sujet donné afin d’adapter de manière dynamique le contenu ou le style d’enseignement » (traduction libre de Murray, 1999, p. 1). Les STI sont conçus pour non seulement enseigner du contenu, mais pour évaluer le progrès d’un étudiant et lui fournir des

rétroactions. On retrouve quelques composantes fonctionnelles principales dans les STI ; le modèle du domaine, le modèle de l'apprenant, le modèle de tutorat, le modèle de rétroaction et le modèle d'interface utilisateur (Slavuj et collab., 2015).

« Le modèle du domaine comprend toutes les connaissances dans un domaine précis qui doivent être enseignées par le système » (traduction libre de Slavuj et collab., 2015, p. 815). Par exemple, pour le STI e-Tutor (Figure 7), développé pour enseigner la langue allemande, le modèle du domaine est constitué du « [...] contenu des trois premiers cours d'allemand universitaires durant lesquels les notions grammaticales de base sont généralement enseignées » (traduction libre de Heift, 2010, p. 444).

Le modèle de l'apprenant suit et supervise l'utilisateur dans ses interactions avec l'outil.

Ce suivi implique la création et la mise à jour constante d'un modèle d'étudiant qui recueille toutes les données nécessaires de l'apprenant, qu'il s'agisse de son parcours d'apprentissage, de la justesse de ses réponses aux questions, du temps qu'il met à proposer une solution ou de sa réussite globale dans la résolution d'une tâche. (Traduction libre de Slavuj et collab., 2015, p. 815)

Le modèle de tutorat comprend tous les règlements qui définissent la manière dont le système devrait fonctionner et « [...] précise quelles actions pédagogiques devraient être entreprises par la suite à chaque activité de l'utilisateur » (traduction libre de Slavuj et collab., 2015, p. 816). Le domaine de rétroaction vise à imiter les pratiques correctives des enseignants. « Les enseignants de langue fournissent souvent des rétroactions aux étudiants concernant leur progrès, leur réussite sur des tâches précises, les aspects linguistiques à travailler, etc. De la même manière, les STI

devraient aussi être capables de répondre adéquatement aux apprenants » (traduction libre de Slavuj et collab., 2015, p. 816).

Dernièrement, le modèle d'interface utilisateur se concentre sur « [...] l'environnement graphique par lequel le contenu éducatif est transmis à l'apprenant » (traduction libre Slavuj et collab., 2015, p. 816). Ce modèle concerne l'aménagement graphique des composantes du système et vise à faciliter les interactions machine-utilisateur (Levy & Caws, 2016).

Les STI semblent comporter plusieurs affordances pour l'apprentissage des langues, notamment une capacité à adapter son style d'enseignement (Murray, 1999) ainsi que le contenu (Heift, 2016). Cela dit, on constate toujours que ces outils ont des limites. D'abord, les STI ont une tendance à trop contraindre le domaine à l'étude horizontalement ainsi que verticalement. On fait référence aux contraintes horizontales quand un système se concentre « [...] sur une seule compétence linguistique (écouter, lire, parler et écrire) ou une seule forme (grammaire, vocabulaire) plutôt que sur une combinaison de ces compétences et de ces formes » (traduction libre de Slavuj et collab., 2015, p. 816). Les contraintes verticales signifient « [...] l'incapacité d'un STI à répondre aux besoins des apprenants avec différents niveaux de compétences linguistiques » (traduction libre de Slavuj et collab., 2015, p. 816). Les contraintes verticales sont surtout en raison du fait que les différents niveaux dans les STI sont souvent fondés sur le niveau de formation (12^e année, 2^e année universitaire, etc.) et ne s'appuient pas sur un cadre de référence. Ainsi, il est difficile pour un apprenant d'évaluer sa compétence linguistique de manière objective pour pouvoir se placer au bon niveau (Slavuj et collab., 2015). Par exemple, un apprenant en onzième année pourrait être plus ou moins avancé dans la langue cible que ce qui figure dans le module de onzième année dans le STI, ce qui signifierait que le STI ne serait pas en train de répondre aux besoins réels de l'apprenant.

L'évaluation des données fournies par un utilisateur (user input en anglais) représente aussi un défi pour les STI. Ces systèmes éprouvent des difficultés avec le traitement et l'évaluation des compétences productives en raison de la liberté d'expression que possède un utilisateur durant les tâches de production écrite et de production orale (Slavuj et collab., 2015). « Des inférences incorrectes concernant les données fournies par l'utilisateur, ses connaissances et ses compétences linguistiques peuvent entraîner une action erronée de la part du système, menant ainsi à un processus d'apprentissage inefficace [...] » (traduction libre de Slavuj et collab., 2015, p. 816).

Les STI ont aussi des limites concernant leurs capacités interactives, surtout pour l'oral. Les STI peuvent ne pas bien saisir le message oral qu'un utilisateur tente de transmettre parce que ces outils n'ont pas été conçus pour tenir compte du paralangage tel que la gestuelle, le langage corporel et de la prosodie (Slavuj et collab., 2015), ce qui pourrait mener à une incompréhension de la part du STI.

2.4 L'IA en enseignement-apprentissage des langues

Les outils d'IA sont déjà intégrés en enseignement-apprentissage des langues sous différentes formes tels que les systèmes de réalité augmentée et de réalité virtuelle et les assistants pédagogiques virtuels. Dans cette section, nous allons nous attarder à quelques outils issus de l'IA et à leur potentiel pour l'apprentissage des langues. Tel que dit plus haut, les STI sont des outils issus de l'IA qui visent à enseigner et à évaluer un utilisateur en adoptant une approche adaptative (Op. cit). Il existe aussi, entre autres, les robots conversationnels (chatbots en anglais), les traducteurs automatiques et les outils d'IA générative (IAG).

2.4.1 Les robots conversationnels

Les robots conversationnels sont des outils informatisés capables de converser avec les êtres humains (Kwon, Shin & Lee, 2023). Ils sont « [...] un type d'IA conversationnelle qui permet des

conversations en langage naturel entre les êtres humains et les machines par écrit ou par oral » (traduction libre de Ji et collab., 2023, p. 49). Les robots conversationnels tels qu'Eliza (1966), Parry (1972), Racter (1983), Jabberwacky (1988) et A.L.I.C.E. (1995) existent depuis longtemps (Belda-Medina & Calvo-Ferrer, 2022). Les premiers robots conversationnels étaient « [...] textuels avec des réponses prédéfinies, basées sur des scénarios de questions-réponses, de sorte que leurs réponses étaient considérées comme prévisibles et que l'interaction n'était pas perçue comme naturelle selon les normes humaines » (traduction libre de Belda-Medina & Calvo-Ferrer, 2022, p. 1). Grâce aux avancées dans les domaines du traitement de langage naturel (TLN) et de la compréhension du langage naturel (CLN), les robots conversationnels ont connu de grandes améliorations (Belda-Medina & Calvo-Ferrer, 2022 ; Ji et collab., 2023 ; Kwon et collab., 2023 ; Ruan et collab., 2021). Conséquemment, les robots conversationnels modernes sont plus sophistiqués que leurs prédécesseurs (Belda-Medina & Calvo-Ferrer, 2022).

Les robots conversationnels sont utilisés en enseignement-apprentissage des langues, surtout pour pratiquer les échanges conversationnels dans une langue cible (Kwon et collab., 2023). Pendant les conversations en L2, les robots conversationnels peuvent créer un environnement d'apprentissage plus détendu pour un apprenant (Belda-Medina & Calvo-Ferrer, 2022 ; Bibauw, François & Desmet, 2022 ; Ji et collab., 2023 ; Kwon et collab., 2023) où il peut développer et organiser ses idées à son propre rythme (Guo, Wang & Chu, 2022). Les robots conversationnels sont aussi très accessibles et ne se fatiguent pas (Belda-Medina & Calvo-Ferrer, 2022 ; Bibauw et collab., 2022). Conséquemment, un apprenant peut répéter et refaire les mêmes exercices plusieurs fois (Belda-Medina & Calvo-Ferrer, 2022 ; Bibauw et collab., 2022 ; Kwon et collab., 2023).

Les principales limites identifiées concernant l'utilisation des robots conversationnels en éducation sont « [...] la nature scénarisée de leurs réponses, ce qui les rend plutôt prévisibles, leur

compréhension limitée (vocabulaire, sens intentionnel) et l'inefficacité de la communication » (traduction libre de Belda-Medina & Calvo-Ferrer, 2022, p. 3). Les robots conversationnels peuvent aussi produire des réponses dénuées de sens (Belda-Medina & Calvo-Ferrer, 2022). Une mauvaise interprétation de la part de la machine pourrait également engendrer une panne de communication (Ji et collab., 2023). Les avancées en IAG semblent adresser la majorité de ces limites.

2.4.2 Les traducteurs automatiques

Un autre type d'outil numérique issu de l'IA est le traducteur automatique. Il y a différents types de traducteurs automatiques qui ont été développés au fil des années. Tout d'abord, il y a les outils qui utilisent la traduction directe, ce qui signifie qu'ils traduisent directement la langue source vers la langue cible sans utiliser d'intermédiaires. Aucune analyse syntaxique ne se produit durant la traduction directe et la traduction se fait mot à mot à partir de données statistiques. Il y a les systèmes de transfert visant à créer des traductions plus idiomatiques grâce à l'intégration d'un système d'analyse syntaxique. Il existe aussi des systèmes basés sur la traduction en passant par une langue intermédiaire ou interlingua (souvent l'anglais) (Poibeau, 2017).

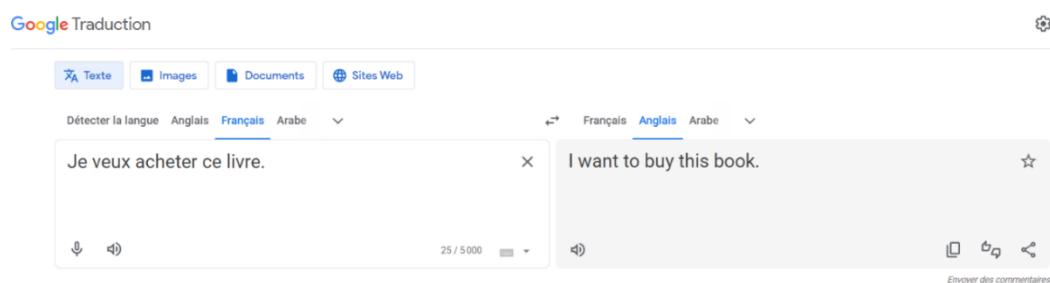


Figure 8 - Le traducteur automatique Google Traduction

L'année 2015 a vu le passage de la traduction automatique statistique (TAS) à la traduction automatique neuronale (TAN) (Barbin, 2020). La principale différence entre la TAN de la TAS

(et des autres systèmes de traduction automatique susmentionnés) est le fait que la TAN est basée sur un système d'autoapprentissage tandis que les autres systèmes se fient aux dictionnaires et aux « [...] modèles statiques construits à partir de corpus découpés en unités de traduction » (Barbin, 2020, p. 51). « Se démarquant clairement de la TAS, la TAN fonctionne sur le modèle du cerveau humain au travers de l'intelligence artificielle. Le saut qualitatif est indéniable : la TAN permet de générer des traductions de meilleure qualité (plus précises et naturelles) qui se rapprochent de la traduction humaine » (Barbin, 2020, p. 51). « [...] [L]a TAN imite le fonctionnement du cerveau humain (création des voies neuronales) pour générer des traductions plus naturelles et idiomatiques » (Barbin, 2020, p. 51), ainsi adoptant une approche semblable à celle utilisée dans l'apprentissage profond avec la création de voies neuronales (Op. cit).

Les traducteurs automatiques sont des outils souvent utilisés par les apprenants de langue (Briggs, 2018 ; Jolley & Maimone, 2015). Bien que l'utilisation de ces outils soit répandue, leur intégration en enseignement-apprentissage des langues est un sujet controversé (Godwin-Jones, 2022). On constate que ces outils comportent des affordances ainsi que des limites en tant qu'outils pédagogiques.

Selon les étudiants recensés dans une étude en 2015, des affordances notables des outils de traduction automatique (TA) sont leur rapidité à fournir une traduction ainsi que leur accessibilité (Fredholm, 2015). En outre, « plusieurs études menées par Fredholm (2014, 2015, 2019) et par E. M. O'Neill (2016, 2019) montrent une amélioration de la qualité de l'écriture grâce à l'utilisation de la TA intégrée dans les tâches d'apprentissage » (traduction libre de Godwin-Jones, 2022).

Les traducteurs automatiques ne sont cependant pas dénués de défauts (Briggs, 2018) et il semble que les étudiants sont au courant de ce fait. Dans l'étude exploratoire de Briggs (2018), la majorité des étudiants recensés ont déclaré n'avoir qu'une confiance limitée en la précision des traductions

par la TA. On constate aussi que les étudiants semblaient rencontrer des difficultés à analyser de manière critique les traductions sollicitées en raison d'un manque de compétences analytiques et critiques qui sont nécessaires pour l'utilisation efficace de ces outils (Briggs, 2018). Enfin, les enseignants semblent être d'accord qu'il faut que les apprenants soient formés à l'utilisation des traducteurs automatiques pour pouvoir les exploiter efficacement (Briggs, 2018 ; Godwin-Jones, 2022 ; Jolley & Maimone, 2015).

2.4.3 L'intelligence artificielle générative

Récemment a vu le jour un type d'outil d'IA relativement nouveau, le générateur de contenu. Cet outil permet la communication et fournit de l'instruction et des rétroactions, mais ce qui le distingue des autres outils, c'est sa capacité de générer du nouveau contenu qui peut être indiscernable du contenu généré par les êtres humains (Rossoni & ChatGPT, 2022 ; Urbain, 2023).

« Les progrès rapides de l'IA générative ont suscité un grand intérêt pour ses applications potentielles en enseignement des langues » (Creely, 2023, p. 1). Il semble que l'IAG comporte plusieurs affordances en enseignement-apprentissage des langues ; ces outils pourraient fournir plus d'occasions pour un apprenant de s'exercer dans la langue cible, suivre le progrès de l'apprenant et fournir des rétroactions synchrones, utiles et personnalisées (Galaczi, 2023). L'interactivité des outils d'IAG a du potentiel pour rendre le processus d'apprentissage plus engageant en offrant aux apprenants du contenu personnalisé et un parcours d'apprentissage plus adapté à leurs besoins individuels (Creely, 2023).

On reconnaît aussi de potentiels défis concernant l'intégration de l'IAG en enseignement-apprentissage des langues, tels que la précision de la production des outils d'IAG (Creely, 2023 ; Galaczi, 2023), c'est-à-dire la qualité de ses réponses. Certains chercheurs craignent aussi que l'utilisation répandue des outils d'IAG puisse nuire à la pensée critique en raison d'une dépendance

excessive à ces outils (Creely, 2023). On parle aussi des limites déontologiques quant à l'intégration de l'IAG en milieu éducatif, surtout en ce qui concerne le respect de la vie privée, la diffusion de ses données personnelles et la paternité du contenu. En raison de ces affordances et de ces défis, Creely (2023) avance qu'il faut donc trouver un équilibre concernant l'intégration des outils d'IAG en enseignement-apprentissage des langues.

Les grandes avancées en TLN (Op. cit) ont rendu les interactions personne-machine de plus en plus sophistiquées, adaptatives et naturelles, ce qui semble proposer des affordances pour l'intégration des outils issus de l'IAG en enseignement-apprentissage des langues (Creely, 2023). Parmi les outils d'IAG, on compte ChatGPT, l'outil auquel la présente étude s'intéresse. Comme on l'a dit plus haut, il existe d'autres modèles comparables à ChatGPT, mais lorsque cette étude a été élaborée, ChatGPT était l'outil d'IAG disponible. D'ailleurs, selon Fang et collab., « ChatGPT est une innovation révolutionnaire dans le domaine des agents conversationnels. Il possède la capacité remarquable de comprendre des instructions complexes et de générer des réponses qui imitent fidèlement celles des êtres humains » [traduction libre] (2023, p. 1).

2.5 Les affordances et les limites de l'IAG pour la RCÉ

Dans cette section, on résume les résultats de plusieurs études portant sur l'utilisation de ChatGPT (et de ses différentes versions) pour des tâches pédagogiques liées à la rédaction et à la révision textuelle afin de mettre d'avant ses affordances pour la RCÉ. On présente ensuite une vue d'ensemble des limites de ces outils identifiées dans les études mentionnées.

2.5.1 La capacité de ChatGPT pour la correction des erreurs grammaticales

À notre connaissance, peu de chercheurs se sont concentrés à ce jour sur l'utilisation de l'IAG pour la RCÉ et encore moins sur la RCÉ interactive. Les recherches qui ont été menées se concentrent surtout sur l'anglais, exception faite de l'étude de Fang et collab. (2023) qui a évalué

la capacité correctrice de ChatGPT pour l'anglais, l'allemand et le chinois. Pendant cette étude, les auteurs ont trouvé que ChatGPT comporte des avantages pour des contextes multilingues et avec peu de ressources (les langues pour lesquelles il y a moins de ressources disponibles), démontrant beaucoup de potentiel pour la correction des erreurs grammaticales (CEG).

D'après les résultats de l'étude préliminaire de Wu et collab. (2023), lorsque comparé à deux autres outils de correction (GECToR et Grammarly), ChatGPT possède une capacité unique d'aller au-delà des corrections ciblées afin de garder la cohérence textuelle sans sacrifier la grammaticalité. ChatGPT a aussi détecté un grand nombre d'erreurs et a conséquemment obtenu le score le plus élevé pour le rappel (Wu et collab., 2023).

Les résultats de l'étude de Fang et collab. (2023) semblent corroborer ceux obtenus par Wu et collab. (2023). Pour la correction des erreurs au niveau de la phrase en anglais, ChatGPT a fait preuve d'une capacité élevée pour le rappel, plus élevée que celles des autres systèmes de pointe. Afin de vérifier les résultats obtenus, une évaluation humaine a été menée pour comparer les corrections de ChatGPT avec celles des autres systèmes de CEG testés ; GECToR et T5 large, ainsi qu'avec celles de Grammarly. Trois étudiants de troisième cycle ont évalué les corrections de ces quatre systèmes. ChatGPT a obtenu les meilleurs scores pour la fluidité et des scores plus bas, mais comparables pour la quantité minimale de modifications et pour la sous-corrrection.

Tel que dit plus haut, l'étude de Fang et collab. (2023) a analysé la CEG en anglais, ainsi qu'en chinois et en allemand. Pour ce faire, ils se sont servis de bases de données contenant des phrases avec des erreurs grammaticales dans ces langues. La base de données utilisée pour l'allemand, Falko-MERLIN³, est considérablement plus petite que celles pour l'anglais et pour le chinois, ce

³ Une base de données avec des phrases contenant des erreurs grammaticales en allemand

qui signifie que la CEG en allemand a pu être considérée une tâche avec peu de ressources. Les auteurs constatent que les résultats de la correction des erreurs grammaticales pour l'allemand et le chinois démontrent des tendances similaires à celles pour la CEG de l'anglais où ChatGPT dépasse les autres systèmes de pointe pour le rappel. Les résultats obtenus en allemand démontrent aussi que ChatGPT peut être utilisé pour la CEG pour des langues avec peu de ressources en plus des langues à haut niveau de ressources.

Fang et collab. (2023) ont aussi analysé la performance de ChatGPT pour la CEG dans le traitement d'un texte entier en anglais en comparant l'outil avec MultiEnc-dec (un système conçu précisément pour la CEG en anglais pour les textes entiers) pour les types d'erreurs à granularité fine basés sur les parties du discours (POS-based fine-grained error types en anglais) dans la base de données BEA19⁴. Les auteurs constatent que ChatGPT démontre une bonne performance dans le traitement des erreurs liées à la ponctuation, aux noms et aux noms possessifs, obtenant des scores considérablement plus élevés pour le critère $F_{0.5}$ (évaluant à la fois la précision et le rappel, mais en accordant plus de poids à la précision) dans ces domaines, lorsque comparé à MultiEnc-dec.

Loem, Kaneko, Takase & Okazaki (2023) ont utilisé JFLEG, CoNLL2014 et W&I+LOCNESS (des bases de données comprenant des phrases avec des erreurs grammaticales en anglais) avec des approches « zero-shot » et « few-shot » avec 16 exemples et ont comparé la performance de GPT-3 (un ancien modèle de ChatGPT) à un modèle Transformer et à la CEG supervisée et non supervisée. Le « zero-shot » se réfère à la capacité d'une machine à correctement résoudre un problème ou à classer un élément inconnu sans avoir été explicitement entraîné pour ce faire (Xian,

⁴ Une base de données avec des phrases contenant des erreurs grammaticales en anglais

Lampert, Schele & Akata, 2020). Chaque « shot » signifie un exemple explicitement partagé avec la machine. Le Tableau 3 résume ces différents contextes. Dans le cadre de l'étude de Loem et ses collègues (2023), la comparaison a été faite selon les scores GLUE (évaluant le minimum de rappel et de précision) pour JFLEG et le score $F_{0.5}$ (Op. cit) pour CoNLL2014 et W&I+LOCNESS. La performance de GPT-3 pour la CEG était comparable à celles des autres systèmes et GPT-3 a dépassé le modèle Transformer pour toutes les bases de données (Loem et collab., 2023).

Type de requête	Définition
Zero-shot	Un contexte dans lequel aucun exemple n'est fourni.
Zero-shot Chain of Thought (CoT)	Un contexte dans lequel aucun exemple n'est fourni, mais on demande explicitement à la machine de penser étape par étape.
Few-shot	Un contexte où l'on partage quelques exemples avec la machine.

Tableau 3 - Tableau de synthèse des contextes de requêtes

2.5.2 L'importance des requêtes

L'étude susmentionnée de Fang et collab. (2023) a évalué non seulement la capacité correctrice de ChatGPT, mais aussi les manières de modifier les requêtes utilisées afin d'interagir avec ce dernier. Les auteurs ont conçu des requêtes pour des contextes « zero-shot » et « zero-shot chain-of-thought » pour utiliser avec ChatGPT pendant le processus de la correction. « Zero-shot chain of thought » (aussi appelé CoT) est une approche qui se base sur l'apprentissage « zero-shot » (Op. cit), mais qui demande explicitement à la machine de raisonner en adoptant une approche d'étape par étape. Kojima et collab. (2023) ont trouvé qu'à travers différents domaines (tâches mathématiques, tâches symboliques, tâches de sens commun et d'autres tâches de raisonnement logique) que l'outil d'IAG dont ils se servaient, GPT-3, obtenait les bonnes réponses simplement grâce à l'ajout de la phrase « pensons étape par étape⁵ » [traduction libre] (p. 1) à la requête initiale.

⁵ « Let's think step by step » en anglais

Fang et collab. (2023) ont testé différentes requêtes avec ChatGPT pendant la correction des erreurs grammaticales en anglais, en allemand et en chinois pour voir en quoi le degré d'explicité des requêtes affectait les réponses obtenues. García-Peñalvo (2023) affirme qu'on obtient deux réponses différentes en posant la même question à ChatGPT dépendamment de si on la pose avec ou sans contexte. Les résultats de l'étude de Fang et collab. (2023) démontrent que la performance de ChatGPT s'améliore considérablement pour la correction des erreurs grammaticales pour deux des cinq bases de données testées en adoptant la démarche CoT pendant la formulation des requêtes. En ajoutant à la requête la phrase « [...] trouvez et corrigez toutes les erreurs grammaticales dans la phrase en gardant la structure de la phrase originale aussi inchangée que possible »⁶ (traduction libre de Fang et collab., 2023, p. 3), les explications des corrections de ChatGPT étaient plus concises et mieux organisées. En utilisant des requêtes sans cette phrase, les réponses obtenues étaient souvent si longues et désorganisées qu'elles nécessitaient de l'intervention manuelle (Fang et collab., 2023).

Les résultats obtenus suggèrent qu'avec les méthodes appropriées de formulation des requêtes, ChatGPT est en mesure de proposer une CEG pour les langues autres que l'anglais, y compris les langues à faibles ressources (Fang et collab., 2023).

D'autres études ont été menées auprès de différentes versions de ChatGPT pour la CEG en anglais. Loem et collab. (2023) se sont aussi intéressés aux manières de modifier les requêtes afin de voir en quoi les réponses de GPT-3 changeaient par la suite. Les auteurs soulignent l'importance de la contrôlabilité pour la CEG pour son application pour les tâches authentiques, surtout en éducation,

⁶ « Identify and correct any grammatical errors in the sentence while keeping the original sentence structure unchanged as much as possible » en anglais

où la capacité à personnaliser les rétroactions selon les niveaux des apprenants et les types d'erreurs précis peuvent affecter l'apprentissage de manière importante (Loem et collab., 2023).

Afin de voir en quoi la manière de modifier la requête affecte la réponse de GPT-3, les chercheurs ont créé trois types de requêtes. « La requête instructive » demandait à GPT-3 de manière explicite de faire la CEG. « La requête trompeuse » demandait une tâche adjacente à la CEG telle que la paraphrase, mais ne demandait pas explicitement une CEG. « La requête impertinente » demandait une tâche non liée à la CEG, telle que la traduction. Les résultats de cette expérience révèlent que les instructions dans les requêtes affectent la performance de GPT-3 pour la CEG de manière importante (Loem et collab., 2023). Loem et collab. (2023) ont aussi employé trois différents types de requêtes précisant les niveaux des rédacteurs des textes soumis afin de voir comment ce paramètre impacte les résultats. Il y avait la requête « information positive » précisant correctement le niveau du rédacteur, la requête « information négative », précisant incorrectement le niveau du rédacteur et la requête « sans information » qui n'a fourni aucune information additionnelle sur le niveau du rédacteur. Parmi ces trois types de requêtes, pour le contexte zero-shot, les requêtes « information positive » ont obtenu les meilleurs résultats tandis que les requêtes « information négative » ont nui aux résultats. Les auteurs affirment que ces résultats soulignent l'importance et l'impact de l'ajout de l'information dans la requête (Loem et collab., 2023). Pour le contexte few-shot, les requêtes « sans information » ont obtenu des résultats comparables à ceux obtenus par les requêtes « information positive », tandis que les requêtes « information négative » ont toujours obtenu les pires scores.

Enfin, les auteurs ont modifié leurs requêtes selon soit l'instruction, soit les exemples afin de déterminer l'élément avec le plus d'impact sur la réponse obtenue. Les auteurs constatent que les exemples semblent être le volet le plus important de la requête, tandis que les instructions peuvent

prendre diverses formes. Cela est probablement parce que les exemples contiennent des informations précises et contextuelles, tandis que les instructions peuvent être plus abstraites (Loem et collab., 2023).

Ainsi, l'étude de Loem et collab. (2023) démontre le rôle important que joue la formulation de la requête pour GPT-3 pour les tâches de CEG en anglais. Ils précisent qu'en créant de manière méticuleuse les requêtes et les exemples, « [...] GPT-3 peut être contrôlé de manière efficace afin de cibler différents aspects pendant la CEG et peut s'adapter aux besoins d'apprentissage uniques et divers » (traduction libre de Loem et collab., 2023, p. 212).

Coyne, Sakaguchi, Galvan-Sosa, Zock & Inui (2023) ont aussi étudié la CEG pour l'anglais, mais avec GPT-3.5 (ChatGPT) et GPT-4 pour les contextes zero-shot et few-shot. Les bases de données utilisées étaient BEA19 et JFLEG et les résultats ont été comparés avec ceux des systèmes de CEG de référence et des corrections humaines selon les scores F_{05} et GLUE (Op. cit). Tout comme Loem et collab. (2023), Coyne et collab. (2023) constatent que l'information contenue dans la requête a un impact significatif sur la performance de l'outil pour les tâches de CEG.

Ces auteurs ont également étudié comment le non-respect des règles relatives à la ponctuation peut impacter la CEG pendant la formulation de la requête. Les résultats de cette expérience démontrent que les deux systèmes (GPT-3.5 et GPT-4) ont rencontré des difficultés à bien effectuer la CEG quand la requête initiale comprenait des erreurs de ponctuation. Cela souligne aussi l'importance de bien formuler les requêtes afin d'obtenir de meilleures réponses.

Pour conclure, Coyne et collab. (2023) affirment que les modèles GPT-3.5 et GPT-4 ont fait preuve d'une bonne performance pour la CEG, ce qui corrobore les résultats obtenus des autres études

susmentionnées. On constate ainsi l'importance de la requête dans le dialogue avec l'IAG pour obtenir une RCÉ.

2.5.3 La traduction automatique avec ChatGPT

Les traducteurs automatiques sont des outils souvent utilisés par les apprenants de L2 (Op. cit) et l'on s'est aussi intéressé à l'utilisation de ChatGPT pour la TA. Jiao et collab. (2023) ont entrepris une évaluation préliminaire sur la capacité de traduction de ChatGPT.

D'abord, ils se sont intéressés à la traduction multilingue et pour ce faire, ChatGPT a été comparé avec trois autres systèmes de traduction automatique ; Google Traduction, DeepL et Tencent TranSmart selon trois métriques pour l'évaluation de la traduction automatique ; le score BLEU, le score ChrF++ et le score TER. Afin d'évaluer la capacité de ChatGPT pour la traduction multilingue, les auteurs ont choisi quatre langues ; l'anglais, l'allemand, le chinois et le roumain. Ces systèmes ont été évalués pour la traduction des phrases provenant de la base de données Flores-101, comprenant 1012 phrases traduites vers 101 langues. Pour formuler leurs requêtes, les auteurs ont sollicité des pistes d'exemples directement de ChatGPT en envoyant la requête : « Fournissez dix requêtes ou modèles concis qui peuvent vous faire traduire »⁷ (traduction libre de Jiao et collab., 2023, p. 1). Les auteurs ont trouvé que ces requêtes ont bien fonctionné et n'avaient que des variations mineures quant à leur efficacité.

ChatGPT a obtenu des scores comparables de ceux obtenus par Google Traduction et par DeepL pour la traduction de l'allemand vers l'anglais ainsi que pour la traduction de l'anglais vers l'allemand. Il semble que ChatGPT fait preuve d'une bonne performance pour la traduction des langues de la même famille (Jiao et collab., 2023).

⁷ « Provide ten concise prompts or templates that can make you translate » en anglais

En évaluant la robustesse des traductions, ChatGPT a obtenu de meilleurs scores que les deux autres systèmes pour la série 3 de la base de données WMT20, qui est constituée de commentaires du site web de médias sociaux Reddit, signifiant qu'elle pourrait contenir des fautes de langue ainsi que du langage familier. Les auteurs postulent que la performance élevée de ChatGPT avec cette base de données est en raison de la nature conversationnelle de l'outil, qui peut conséquemment mieux générer du texte qui ressemble à la langue parlée (Jiao et collab., 2023).

Afin d'explorer d'autres manières de solliciter une traduction, la requête « pivot » a été testée aussi. Cette méthode demande à l'outil d'utiliser une langue intermédiaire afin de traduire un segment d'une langue source vers une langue cible. Cette stratégie est similaire à celle employée par les traducteurs automatiques employant la traduction par interlingua (Op. cit). Dans le cadre de l'étude de Jiao et collab. (2023), la langue intermédiaire utilisée a été l'anglais en raison de son statut comme langue à haut niveau de ressources. Les requêtes pivot ont été utilisées pour les langues distantes, c'est-à-dire des langues qui appartiennent à différentes familles de langues. Un exemple d'une requête pivot citée dans l'étude est la suivante : « Veuillez fournir la traduction anglaise d'abord, puis la traduction chinoise pour ces phrases, une par une : »⁸ (traduction libre de Jiao et collab., 2023, p. 5) utilisée pour traduire un texte en allemand en chinois en passant par l'anglais. Les auteurs constatent qu'en employant cette technique, les traductions de ChatGPT ont été améliorées pour les langues distantes de façon notable selon les essais auprès de la version 2023.01.31 de l'outil (Jiao et collab., 2023).

Avec les nouvelles versions de ChatGPT, sa capacité pour la traduction s'est améliorée de manière significative. GPT-4 (précisément la version 2023.03.15) a obtenu de meilleurs scores lorsque

⁸ « Please provide the English translation first and then the Chinese translation for these sentences one by one: » en anglais

comparé avec l'ancienne version testée. Les nouveaux scores plus élevés étaient comparables avec ceux des autres systèmes de traduction automatique de pointe. Les auteurs rendent aussi explicite le fait que leurs essais avec ChatGPT étaient dans un contexte zero-shot et ils postulent qu'avec plus de contexte (telle que la stratégie CoT), les traductions pourraient être améliorées encore (Jiao et collab., 2023).

Enfin, on a employé une dernière métrique, l'évaluation humaine, afin d'évaluer les traductions par ChatGPT. On a demandé à trois annotateurs de repérer les erreurs dans les traductions fournies par ChatGPT, GPT-4 et Google Traduction. Les annotateurs ont fait une évaluation manuelle afin de repérer les cas de sous-correction, de surcorrection ainsi que les erreurs de traduction. GPT-4 a obtenu les scores les plus bas pour ces trois types d'erreurs, ce qui corrobore le fait que GPT-4 est un outil de traduction puissant (Jiao et collab., 2023).

2.5.4 La rédaction assistée par l'IAG

Une étude a été menée par Su, Lin & Lai (2023) afin d'explorer les affordances et les limites de ChatGPT pendant les différentes étapes de la rédaction d'un texte argumentatif en anglais. Les auteurs se sont servis de ChatGPT pour la rédaction d'un plan pour un texte argumentatif dans le cadre d'une tâche de rédaction simulée. Les auteurs ont trouvé que pendant l'étape de la planification, ChatGPT pouvait suggérer des pistes d'amélioration concernant le plan pour un texte argumentatif en analysant les liens entre les affirmations et les sous-affirmations, en évaluant la qualité des preuves et des réfutations proposées et en suggérant des contre-arguments, ce qui pourrait améliorer l'argumentation sur le plan structurel (Su et collab., 2023). Cela dit, les auteurs affirment que la qualité des suggestions proposées par ChatGPT dépendait de l'information partagée, comme les rubriques pour évaluer la production écrite.

Pendant l'édition du texte, les rétroactions proposées par ChatGPT ont bien identifié les forces et les faiblesses dans le texte partagé et ont proposé des pistes d'amélioration, mais elles étaient vagues et abstraites (Su et collab., 2023). Néanmoins, les auteurs affirment qu'un utilisateur peut demander des précisions additionnelles sur les suggestions proposées par l'outil. Afin de démontrer qu'il est possible de demander des précisions sur les suggestions proposées par ChatGPT, les auteurs lui ont posé une question de suivi et ensuite ChatGPT a fourni une liste de réponses plus détaillées. « La capacité de ChatGPT de générer des réponses aux questions posées par les utilisateurs lui permet d'échafauder l'aspect dialogique de la rédaction argumentative en créant un espace de conversation à faible enjeu où les utilisateurs peuvent explorer différentes perspectives sur le sujet donné » (traduction libre de Su et collab., 2023, p. 8).

Concernant la révision linguistique et la correction grammaticale, ChatGPT s'est avéré utile selon ces auteurs. Pour la correction au niveau de la phrase, ChatGPT a été en mesure de juger la précision grammaticale, de développer le sens qu'une phrase tente de transmettre et de proposer des reformulations. Pour la correction d'un texte entier, ChatGPT a corrigé les erreurs grammaticales (telles que les erreurs concernant les déterminants, les verbes de modalité, les temps verbaux ainsi que les erreurs de frappe) et a modifié le choix de mots et les structures syntaxiques pour assurer que les formulations étaient précises, ainsi que pour améliorer la concision et le ton professionnel du texte (Su et collab., 2023). Les auteurs soulignent que c'est la responsabilité de l'utilisateur d'exercer son esprit critique afin de juger si les rétroactions proposées par ChatGPT devraient être acceptées ou rejetées. Cependant, Su et collab. (2023) affirment qu'il est possible d'entrer en interaction avec l'outil afin d'obtenir des précisions. Ils suggèrent aussi que ChatGPT peut être utilisé comme moteur de recherche pour trouver des collocations, pour discerner les mots similaires et afin de vérifier les usages grammaticaux (Su et collab., 2023).

Enfin, comme ChatGPT sauvegarde les fils de discussion, les auteurs postulent que cela pourrait faciliter une réflexion sur le processus de rédaction après la réalisation de la tâche. À partir des résultats obtenus pendant cette étude exploratoire, les auteurs affirment que ChatGPT peut fournir des réponses et des commentaires personnalisés, évaluer et faire des suggestions sur l'organisation du contenu, effectuer une analyse linguistique et faire une relecture du texte (Su et collab., 2023).

2.5.5 Les limites de l'IAG

Bien que l'IAG comporte des affordances, dont on vient de faire l'état, elle comporte aussi des limites. Notamment, il y a le fait qu'elle ne produit pas nécessairement des réponses correctes. En raison de l'algorithme de prédiction textuelle que suit ChatGPT, le système tente de fournir la réponse la plus probable (OpenAI, 2024c). Conséquemment, ChatGPT peut produire des réponses erronées (on en a fait beaucoup état dans les journaux et les médias). ChatGPT peut aussi partager des informations « fabulées » (Bang, 2023 ; Kaplan-Rakowski, Grotewold, Hartwick & Papin, 2023). Les « fabulations » font référence aux occasions où ChatGPT fournit des informations incorrectes ou qui ne peuvent pas être vérifiées, les présentant comme des informations factuelles. On parle aussi de limites déontologiques quant à l'utilisation de ChatGPT, surtout en ce qui concerne l'intégrité académique et la paternité du contenu (Su et collab., 2023).

L'étude préliminaire de Wu et collab. (2023) démontre que ChatGPT a une tendance à surcorriger. Un exemple cité dans leur article est une occasion où ChatGPT a remplacé le mot « chances » par « opportunités » dans un texte anglais, bien que le terme original fût acceptable en contexte. Ainsi, cette révision pourrait être considérée comme étant superflue. Ces auteurs postulent que c'est peut-être en raison de la capacité de ChatGPT à générer un même texte de diverses manières en tant que GML. L'étude menée par Fang et collab. (2023) a obtenu des résultats similaires. Certains

étudiants pourraient trouver accablante la quantité de corrections fournies par ChatGPT lors d'une RCÉ exhaustive et seraient conséquemment découragés. Les résultats d'une enquête portant sur la quantité de RCÉ idéale ont trouvé que certains étudiants ont décrit la RCÉ comprehensive comme étant accablante ou déprimante (McMartin-Miller, 2014).

D'autre part, ChatGPT semble éprouver des difficultés avec certains types d'erreurs pendant la CEG telles que les accords, la coréférence, la concordance des temps et les erreurs de délimitation à travers les phrases (Fang et collab., 2023), mais il est important de noter que ces données reflètent la CEG pour l'anglais, l'allemand et le chinois. Ces auteurs constatent aussi que la performance de ChatGPT a été plus faible en traitant les erreurs s'étendant au-delà de la délimitation de phrases, comme démontré par un score F_{05} considérablement plus faible pour les erreurs de conjugaison verbale et de ponctuation. MultiEnc-dec a dépassé ChatGPT pour la correction des erreurs liées aux accords sujet-verbe, aux prépositions, aux noms numéraux, aux déterminants et aux pronoms pour les textes entiers. Les auteurs postulent que la sous-performance de ChatGPT pourrait être attribuée à des limitations potentielles concernant sa mémoire contextuelle, la consistance et la cohérence entre les phrases dans un document entier (Fang et collab., 2023).

En ce qui concerne la traduction, Google Traduction et DeepL ont dépassé ChatGPT pour la traduction du roumain vers l'anglais et de l'anglais vers le roumain (Jiao et collab., 2023). On constate aussi que la performance de ChatGPT a été plus faible pour la paire de langues constituée du chinois et de l'allemand comparée à sa performance pour la paire de langues comprenant l'anglais et l'allemand (Jiao et collab., 2023). L'anglais et l'allemand sont des langues dans la même famille (les langues germaniques) et utilisent tous les deux l'alphabet latin, tandis que le chinois (mandarin) est une langue chinoise qui utilise des sinogrammes. Les auteurs postulent que ChatGPT a proposé de meilleures traductions pour les langues appartenant à la même famille en

raison d'un transfert de connaissance de meilleure qualité lorsqu'on reste dans une seule famille de langues (Jiao et collab., 2023). Cela dit, les autres systèmes de traduction automatique ont aussi dépassé la performance de ChatGPT pour la traduction de l'anglais et l'allemand. Les auteurs ont aussi noté que l'écart entre la performance de ChatGPT et celle des autres systèmes devient plus grand lorsqu'on utilise une paire de langues dont les deux langues ont le statut de langues à faible niveau de ressources, telles que le chinois et le roumain (Jiao et collab., 2023). Enfin, pour la robustesse de la traduction, ChatGPT n'a pas performé de manière comparable à Google Traduction et DeepL pour deux des trois bases de données utilisées.

En ce qui a trait à la rédaction, Su et collab., affirment que bien que ChatGPT semble être en mesure d'échafauder le processus de la rédaction, cet échafaudage est limité parce que l'outil ne peut pas entamer un dialogue et ne génère du contenu qu'après la réception d'une requête de l'utilisateur. Cela signifie que ChatGPT ne peut pas poser ses propres questions afin de mieux appuyer l'utilisateur ; c'est un modèle réactif. Ainsi, ChatGPT est limité dans sa capacité à fournir des conseils proactifs sur l'écriture.

Les auteurs constatent également que dans la liste d'arguments proposés par ChatGPT pour répondre à la question de suivi posée pendant l'étape de la planification, le contenu se répétait et se chevauchait, présentant des redondances (Su et collab., 2023). Enfin, pendant l'évaluation du contenu, les rétroactions fournies par ChatGPT étaient surtout de nature évaluative plutôt que des rétroactions explicatives. Les auteurs ont aussi découvert que ChatGPT fournissait des rétroactions différentes après chaque soumission du texte, ce qui remet en question leur fiabilité. Ils avancent que ce manque de consistance dans les rétroactions limite leur utilité (Su et collab., 2023). Par conséquent, les auteurs ont conclu que ChatGPT n'était pas très utile pour l'évaluation du contenu dans le cadre de leur étude.

2.6 Les perceptions de l'IAG en milieu éducatif

Après avoir discuté des affordances et des limites des outils d'IAG, il importe également de mentionner la manière dont les enseignants et les apprenants les envisagent en milieu éducatif. On mentionne d'abord les perceptions des enseignants envers l'IAG et par la suite celles des apprenants. Il est important de s'intéresser aux attitudes des enseignants parce que « si la phase initiale d'examen des attitudes des enseignants à l'égard de la technologie issue de l'IAG est omise, l'intégration de cette technologie dans l'enseignement risque d'échouer » (traduction libre de Kaplan-Rakowski et collab., 2023, p. 320).

2.6.1 Les perceptions des enseignants

Kaplan-Rakowski, Grotewold, Hartwick & Papin (2023) proposent une étude quantitative portant sur les perceptions des enseignants sur l'utilisation, l'intégration et le potentiel de l'IAG en éducation, ainsi que leurs principales inquiétudes.

Un court sondage inspiré du sondage validé de Wozney, Venkatesh & Abrami (2006) a été créé avec un critère d'inclusion précisant que les participants devaient être des enseignants ayant utilisé l'IAG au moins une fois. 147 participants ont répondu à ce critère. Les enseignants sondés avaient des opinions majoritairement positives envers l'IAG en éducation. Ils ont exprimé des sentiments d'optimisme et ont commenté les affordances de l'IAG, telles que « [...] l'amélioration des possibilités pour l'autoapprentissage [...] » (traduction libre de Kaplan-Rakowski et collab., 2023, p. 327). Les chercheurs ont aussi exploré la relation entre les approches et les méthodes employées par les enseignants ainsi que leurs perspectives envers l'IAG. Les résultats ont démontré que les enseignants avaient des perceptions positives envers l'IAG, peu importe leur style d'enseignement ou leur approche pédagogique. 54 % des enseignants sondés ont répondu qu'ils étaient soit susceptibles ou très susceptibles d'intégrer l'IA dans leur enseignement. En outre, « [...] plus de

80 % des enseignants sondés envisageant ChatGPT comme un outil pédagogique précieux pour les étudiants » (traduction libre de Kaplan-Rakowski et collab., 2023, p. 329). Les résultats de cette étude démontrent que plus les enseignants sont au courant de l'IAG, plus ils l'utilisent.

Une question ouverte à la fin du sondage visant à récolter des données portant sur les affordances et les limites perçues de l'intégration de l'IAG en éducation n'a obtenu que 18 réponses, cependant, les sentiments exprimés dans ces réponses corroborent les résultats obtenus à partir des données quantitatives. En ce qui concerne les affordances identifiées par les participants, on retrouve la facilité d'utiliser l'IAG et sa capacité à permettre de gagner du temps pour non seulement les apprenants, mais les enseignants aussi.

Les principaux défis ou limites identifiés sont la crainte que l'utilisation de ChatGPT puisse inciter les apprenants à la paresse (s'ils utilisent ChatGPT pour effectuer leurs travaux sans s'y impliquer) et que l'utilisation de ChatGPT puisse nuire au développement de la pensée critique (Kaplan-Rakowski et collab., 2023).

2.6.2 Les perceptions des apprenants

Shoufan (2023) a mené une étude en deux parties afin d'analyser les perceptions de l'IAG des étudiants. Cette étude avait pour objectif d'explorer les perceptions des étudiants à l'égard de ChatGPT en contexte éducatif et non sa capacité corrective. Les participants étaient 56 étudiants inscrits à un programme d'ingénierie informatique. D'abord, les participants ont effectué une tâche pédagogique (un test formatif portant sur les questions conceptuelles et les tâches de codage) en utilisant ChatGPT. Ils ont ensuite évalué l'outil en répondant à une question ouverte. Les réponses des participants ont été analysées afin de faire ressortir les codes et les thèmes saillants. Ensuite, ces codes et thèmes ont été utilisés afin de créer un questionnaire auquel les participants ont répondu après avoir complété trois autres tests formatifs (semblables au premier) à l'aide de

ChatGPT. Les résultats de cette étude démontrent que les étudiants sondés avaient des perceptions de ChatGPT plutôt positives, exprimant de l'intérêt, de l'admiration, de la motivation et de l'optimisme pour l'outil (Shoufan, 2023). Il y avait quelques étudiants exprimant des sentiments d'incertitude envers ChatGPT et les manières dont il pourrait influencer sur leurs vies. Le principal problème ou défi rapporté a été le fait que ChatGPT ne remplace pas l'intelligence humaine (Shoufan, 2023).

En ce qui concerne l'interaction avec ChatGPT, la majorité des étudiants ont trouvé l'outil facile à utiliser, tandis que la création des requêtes a été modérément difficile, mais le fait de poser des questions additionnelles pouvait aider à trouver la réponse correcte. La justesse des réponses de ChatGPT a obtenu des scores modiques et la majorité des étudiants ont répondu qu'ils croient que l'outil a besoin d'être amélioré. Finalement, la majorité des étudiants ont trouvé ChatGPT utile pour l'apprentissage (Shoufan, 2023).

2.7 Synthèse critique

Tel que dit plus haut, la RCÉ est importante au développement de la production écrite en L2 (Woodworth & Barkaoui, 2020) et les enseignants sont majoritairement d'accord sur cette importance (Bitchener & Ferris, 2011 ; Hamel & Bibeau, 2021). On constate également que la RCÉ représente une tâche exigeante pour les enseignants (Hamel & Bibeau, 2021 ; Lim & Phua, 2019). Des études tentent de démontrer que la RCÉ des enseignants peut comporter plusieurs défauts, telles qu'un manque de cohérence, de clarté et de précision (Woodworth & Barkaoui, 2020). La RCÉ des enseignants peut aussi être différée, bien que la RCÉ synchrone puisse s'avérer plus utile (Arroyo & Yilmaz, 2018 ; Shintani, 2016 ; Shintani & Aubrey, 2016).

Différents outils numériques capables de fournir une RCÉ ont été conçus, tels que les correcticiels et les SÉA. Ces outils comportent plusieurs affordances, notamment une RCÉ quasi immédiate et

l'option de resoumettre un travail plusieurs fois (Zhang & Hyland, 2018), mais ils comportent cependant des défauts importants (Op. cit). Les études démontrent que les correcticiels ne signalent pas toujours toutes les erreurs dans un texte (López Santiago, 2016 ; Ouellet, 2013) et qu'il faut que l'utilisateur soit en mesure d'évaluer la RCÉ proposée pour juger sa pertinence (Hebbinckuys et collab., 2021 ; Ouellet, 2013). On constate que les SÉA sont des systèmes peu adaptables (Godwin-Jones, 2022 ; Woodworth & Barkaoui, 2020) et que leur RCÉA peut être générique (Ranalli, 2018). Cette nature générique permet peu (voire pas) d'interaction ou de négociation avec l'utilisateur (Ranalli, 2018).

Koltovskaia (2020) a mené une étude de cas naturaliste auprès de deux apprenants universitaires d'anglais langue seconde portant sur leur engagement avec la RCÉA de Grammarly pour la rédaction d'une revue de la littérature. En 2020, Grammarly était plutôt un outil pour la révision qu'un outil pour la rédaction (comme il l'est aujourd'hui grâce aux avancées de l'IAG). Pendant cette étude, un apprenant avait rejeté une correction proposée, la croyant fautive, alors qu'elle était juste. Pendant cette même étude, l'autre participante avait accepté toutes les rétroactions proposées par Grammarly portant sur l'emploi des déterminants, exception faite d'une seule rétroaction, peu importe si ces dernières étaient correctes ou incorrectes. Pendant une entrevue, l'étudiante a dit que si elle n'est pas certaine, elle accepte les rétroactions (Koltovskaia, 2020). Sachant ces comportements possibles, il est important pour les apprenants de développer un esprit critique vis-à-vis de la RCÉ générée automatiquement, de bien la comprendre pour être en mesure de juger de sa pertinence afin de l'accepter ou le rejeter, le cas échéant (Hebbinckuys et collab., 2021 ; Koltovskaia, 2020). Ainsi, il pourrait être bénéfique d'interagir avec la machine fournissant la RCÉ. Bien que les STI soient des systèmes plus adaptatifs visant à instruire et à corriger, ils

peuvent être trop contraignants et peuvent aussi être limités dans leurs capacités communicatives (Slavuj et collab., 2015). À notre avis, ces systèmes manquent d'interactivité.

On a aussi dit que pour que la RCÉ soit utile, il faut que les apprenants soient activement engagés avec cette dernière (Op. cit). Selon Zhang, « [...] il est peu probable que les étudiants utilisent les commentaires de la RCÉA qui les confondent et les démoralisent » [traduction libre] (2017, p. 325). Ainsi, il apparaît qu'il serait bénéfique pour les apprenants d'entrer en interaction avec l'outil fournissant la RCÉ afin de mieux la comprendre, en écho à ce qu'a dit Zhang à propos de l'engagement efficace avec la RCÉ (Op. cit). Les capacités conversationnelles et interactives de ChatGPT nous amènent à penser qu'un apprenant pourrait demander à l'outil de fournir des explications personnalisées (Su et collab., 2023) et de répondre aux questions de suivi (Jiao et collab., 2023) pour évaluer la justesse d'une révision proposée. Comme démontré par Koltovskaia (2020), la RCÉA peut être trop générique pour être utile.

D'ailleurs, ChatGPT semble offrir plusieurs affordances propices à la RCÉ pour les apprenants de L2, telles qu'une capacité pour la CEG (Coyne et collab., 2023 ; Fang et collab., 2023 ; Wu et collab., 2023), une utilité pour les contextes multilingues (Fang et collab., 2023; Jiao et collab., 2023), ainsi que sa nature interactive (OpenAI, 2024b ; Su et collab., 2023). Sa performance pendant les tâches de CEG et son utilité pour les contextes multilingues nous amènent à penser que ChatGPT serait en mesure de détecter les erreurs commises fréquemment par les apprenants de FLS (Op. cit) et de proposer une RCÉ utile afin d'y remédier. En outre, sa nature interactive pourrait rendre plus utile la RCÉ. Les apprenants peuvent entrer en conversation avec ChatGPT pour poser des questions de suivi (Jiao et collab., 2023) et pour solliciter une RCÉ personnalisée (Su et collab., 2023).

Ainsi, nous postulons que ChatGPT possède la capacité de non seulement corriger le texte d'un apprenant de FLS (Coyne et collab., 2023 ; Fang et collab., 2023 ; Wu et collab., 2023), mais aussi de lui indiquer comment et pourquoi il faut le corriger (Su et collab., 2023).

2.8 Questions de recherche

Sachant la recension des écrits ci-dessus, on constate que les études sur l'usage des outils d'IA en contexte d'enseignement et d'apprentissage des L2 ne commencent qu'à émerger et se concentrent principalement sur l'anglais. Nous souhaitons ainsi par cette recherche contribuer au domaine en nous concentrant de manière plus précise sur les affordances de ChatGPT pour la RCÉ interactive en contexte d'apprentissage du FLS. Nous tenterons de répondre aux questions suivantes :

QR1. Quelles sont les pratiques autocorrectives actuelles déclarées par les étudiants dans le cours FLS 2751 concernant la révision textuelle des tâches de rédaction?

QR2. Comment est-ce que les étudiants dans le cours FLS 2751 se sont servis de ChatGPT pour obtenir de la RCÉ afin de réviser leur texte écrit?

QR3. Après l'intervention, est-ce que les attitudes envers l'IA des étudiants en FLS 2751 ont changé?

3 Méthodologie

Cette section présente d'abord le cadre théorique dans lequel se situe la présente étude afin de démontrer qu'elle se repose sur des bases théoriques solides. Ensuite, on contextualise et explique l'intervention menée afin de faire la collecte de données. Enfin, on présente les manières dont les données récoltées ont été analysées.

3.1 Cadre théorique

La nature interactive et adaptative de ChatGPT (Su et collab., 2023) pourrait accorder aux apprenants plus d'agentivité et d'autonomie au sein du processus de la révision textuelle et conséquemment promouvoir l'engagement, ce qui rend la RCÉ sollicitée plus utile (Zhang, 2017). Cette thèse de maîtrise vise à explorer l'interactivité du processus de la révision textuelle avec ChatGPT ainsi que l'engagement des apprenants avec l'outil. Étant donné ce contexte, ces recherches sont appuyées par la théorie de l'engagement des apprenants avec la tâche en apprentissage des langues (Egbert, 2020) et par une perspective interactionniste issue de la théorie socioculturelle (Vygotsky, 1987). L'agentivité et l'autonomie promeuvent l'engagement (Mercer, 2019) et l'on constate qu'il existe un lien entre l'engagement des étudiants et de nombreux résultats d'apprentissage désirables tels qu'une meilleure réussite scolaire, la motivation, l'auto-efficacité (Cf. Christenson & Reschly, 2022 ; Mercer, 2019) ainsi que l'utilisation plus efficace de la RCÉ (Op. cit).

Différentes définitions ont été formulées pour l'agentivité et nous retenons celle d'Oga-Baldwin où l'agentivité signifie « [...] les changements apportés par un apprenant pour répondre à ses propres besoins [...] » [traduction libre] (2019, p. 4). D'autre part, selon Bandura, « les principales caractéristiques de l'agentivité permettent aux personnes de jouer un rôle dans leur développement personnel, leur adaptation et leur renouvellement en fonction de l'évolution de la situation »

[traduction libre] (2001, p. 2). L'agentivité et l'autonomie sont des concepts interreliés ; « l'autonomie désigne le besoin pour les apprenants de disposer d'une agentivité et d'être en mesure de gérer leurs propres expériences et actions d'apprentissage » (traduction libre de Mercer, 2019, p. 648). La nature interactive et adaptative de ChatGPT (Op. cit) semble soutenir l'agentivité et l'autonomie des apprenants, ainsi augmentant leur degré d'engagement pendant la révision textuelle.

La théorie de l'engagement avec la tâche linguistique est axée sur les résultats de l'engagement pendant l'apprentissage plutôt que sur les concepts appris en tant que tels (Egbert, 2020). On constate que l'engagement est un élément crucial en apprentissage des langues (Mercer, 2019 ; Oga-Baldwin, 2019). Il y a différents volets de cette théorie tels que l'engagement behavioriste, l'engagement émotionnel, l'engagement cognitif et l'engagement social (Nakamura, Phung & Reinders, 2021). Selon Nakamura, Phung et Reinders (2021), l'engagement behavioriste fait référence au temps dépensé sur une tâche. Ces auteurs précisent que l'engagement émotionnel peut être défini selon différents construits affectifs, comme la participation motivée (Cf. Skinner, Kindermann & Furrer, 2009 ; Skinner, Kindermann, Connell & Wellborn 2009), l'intentionnalité et l'autonomie (Cf. Baralt et collab., 2016) et les émotions positives qui facilitent la réalisation de la tâche ainsi que les émotions négatives qui nuisent à la réalisation de la tâche (Cf. Reeve, 2012). L'engagement cognitif signifie l'attention soutenue et la vigilance pendant la réalisation d'une tâche (Cf. Svalberg, 2009 ; Nakamura et collab., 2021). Dans le contexte des interactions avec les pairs, les indicateurs verbaux tels que la négociation de sens et de forme ainsi que l'autocorrection peuvent être des indicateurs d'engagement cognitif et sont souvent adoptés pour l'analyse de l'enseignement basé sur les tâches (Nakamura et collab., 2021). L'engagement social peut être défini comme le degré de participation d'un individu au sein d'une interaction (Nakamura et

collab., 2021). Étant donné que cette étude exploratoire s'attarde sur l'interaction entre l'apprenant et l'outil (ChatGPT), on s'intéresse surtout à l'engagement social et à l'engagement cognitif.

Bitchener & Ferris (2011) constatent que le développement linguistique, comme exemple du développement cognitif, est facilité par le biais des interactions sociales ; conséquemment, quand les apprenants collaborent et interagissent avec des personnes plus avancées ou informées, leurs compétences linguistiques se développent. Conséquemment, dans le cadre de cette étude, on adopte aussi une perspective interactionniste. La perspective interactionniste est issue de la théorie socioculturelle (Vygotsky, 1987). En apprentissage des langues, la théorie socioculturelle est une approche dialogique, c'est-à-dire, orientée par le dialogue et l'interaction. La plupart des chercheurs en acquisition des langues secondes reconnaissent le rôle de l'interaction, en particulier, le rôle de la négociation du sens qui se passe pendant une interaction, comme une condition nécessaire pour l'acquisition d'une langue seconde (Johnson, 2008). Selon Johnson (2008), les dialogues collaboratifs sont des outils pour la construction des savoirs. En interagissant et en étant engagés activement dans un dialogue collaboratif, les apprenants ont l'occasion de découvrir non seulement ce qu'ils peuvent faire avec la langue, mais aussi ce qu'ils ne peuvent pas faire (Johnson, 2008).

3.2 Intervention et collecte de données

Une collecte de données a été menée auprès de deux groupes d'étudiants du cours *FLS 2751 – Discutons en français* à l'Université d'Ottawa (plan de cours en Annexe 1). Les étudiants inscrits au cours sont de niveau intermédiaire faible. Avant de mener la collecte de données, on a obtenu le consentement des participants par le biais d'un formulaire de consentement partagé sur le site web du cours. La collecte de données auprès du premier groupe d'étudiants a été effectuée pendant

un cours en ligne et celle auprès du deuxième groupe a été menée pendant un cours hybride avec la majorité des étudiants en présentiel.

La collecte de données a été effectuée en trois étapes ; un questionnaire préintervention portant sur les pratiques autocorrectives des participants, une intervention ponctuelle menée auprès de deux sections du cours FLS 2751 pendant laquelle les étudiants ont interagi avec ChatGPT pour solliciter une RCÉ et un questionnaire post-intervention portant sur cette expérience. Les deux questionnaires ont été créés sur la plateforme Microsoft Forms et la professeure a partagé les hyperliens vers ces questionnaires sur le site web du cours.

Le cours FLS 2751, offert par la professeure Martine Rhéaume à l'Université d'Ottawa pendant le trimestre d'automne de 2023, a été choisi pour plusieurs raisons. Un des résultats d'apprentissage attendus est qu'à la fin du cours, l'étudiant soit capable « [d']effectuer une rétroaction sur un énoncé pour le renforcer, l'atténuer, le rectifier, le clarifier, etc. » (Rhéaume, 2023, p. 2). La rétroaction est donc une composante importante au cours. De plus, parmi les aspects étudiés dans le cours, on retrouve le vocabulaire, les techniques d'extension et les techniques de reformulation (Rhéaume, 2023), des aspects qui nous apparaissent comme des éléments linguistiques qui pourraient être travaillés avec ChatGPT. Les apprenants dans le cours FLS 2751 sont de niveau intermédiaire faible ; ce sont des utilisateurs du français indépendants, mais non avancés. Ils possèdent donc assez de connaissances pour pouvoir s'exprimer en français, sans maîtriser la langue, ce qui laisse de la place à la rétroaction.

Dans son cours, la professeure propose une tâche appelée le « Vox Pop » (consignes en Annexe 2) pour laquelle les apprenants doivent choisir un sujet et ensuite créer des questions associées afin de mener des entrevues enregistrées. Les apprenants doivent mener leurs Vox Pop en situation authentique sur campus ou ailleurs dans la ville. La tâche consiste en la création d'une vidéo qui

comporte une introduction, des entrevues avec des participants réels et finalement une conclusion synthétisant les réponses de ces derniers. Le Vox Pop est la première évaluation sommative du cours et vaut 15 % de la note. En raison du calendrier du cours et des contraintes temporelles, les étudiants avaient déjà effectué ce travail et étaient familiers avec cette tâche quand la collecte de données a été menée. Dans le cadre de l'intervention, on leur a demandé de créer une introduction (100 à 150 mots) pour un deuxième Vox Pop hypothétique (une tâche qui n'a pas été menée à bien). Tous les étudiants ont participé à l'intervention, toutefois, seules les données des étudiants ayant fourni leur consentement ont été considérées pour cette étude.

Les étudiants ont d'abord répondu au premier questionnaire (en Annexe 3) qui portait sur leurs pratiques de révision textuelle et d'autocorrection actuelles, ainsi que leurs opinions sur l'IA.

L'intervention ponctuelle s'est déroulée sur une période de classe d'une heure et vingt minutes et a été animée par l'auteure de cette thèse. ChatGPT a été présenté à la classe et la tâche a été expliquée. Les étapes à suivre pour réviser son texte avec ChatGPT ont été énumérées dans la présentation (en Annexe 4). La démarche pour solliciter une RCÉ était :

- Ouvrez une session avec ChatGPT et écrivez la requête « révisé ce texte : » suivie de l'introduction de votre nouveau Vox Pop. Vous écrivez cette introduction directement dans ChatGPT. L'introduction sera de 100 à 150 mots.
- Vous recevrez ensuite votre texte révisé.
- Écrivez la requête « explique toutes les révisions ».
- Assurez-vous de comprendre toutes les explications. Si vous ne comprenez pas les corrections, demandez des précisions ou des explications supplémentaires. C'est à vous de décider comment formuler ces requêtes. Demandez autant de précisions que nécessaire.

- Si vous croyez que ChatGPT s'est trompé, entrez en conversation afin de négocier le sens. C'est à vous de décider comment formuler ces requêtes. Négociez autant que nécessaire.

Une interaction avec ChatGPT a été modélisée afin de démontrer comment s'en servir pour la révision textuelle. Quelques exemples de requêtes à utiliser pour solliciter une rétroaction ont aussi été partagés. Toutefois, on a laissé de la place aux apprenants pour choisir la manière d'échanger avec ChatGPT. Conséquemment, la manière dont ils ont interagi avec ChatGPT a été de leur propre initiative ; c'est ultimement ce qu'on souhaitait : que les interactions entre les étudiants et ChatGPT soient aussi authentiques que possible. La présentation de ChatGPT, l'explication de la tâche et la modélisation d'une interaction ont duré environ vingt minutes. Les étudiants avaient conséquemment une heure pour rédiger et réviser leurs textes.

Pendant l'intervention, les étudiants ont d'abord rédigé leur introduction directement dans ChatGPT pour éviter qu'ils reçoivent au préalable des rétroactions automatiques d'un logiciel de traitement de texte. Ils ont ensuite sollicité une RCÉ de ChatGPT en suivant la démarche présentée. La chercheuse (l'auteure de cette thèse) était présente, mais n'est pas intervenue pendant cette étape afin de préserver l'authenticité des données. Pendant que les étudiants interagissaient avec ChatGPT, elle a effectué d'autres travaux non reliés à la collecte de données. Les fils de discussion sont automatiquement sauvegardés par ChatGPT, donc une fois les interactions terminées, les étudiants ayant accepté de participer à cette étude ont envoyé les hyperliens vers leurs conversations par courriel.

Suite à l'intervention, les étudiants ont rempli le questionnaire post-intervention (en Annexe 5) portant sur cette expérience et sur leurs attitudes (changées ou non) envers l'IA en contexte de révision textuelle. Les réponses aux deux questionnaires ont été automatiquement enregistrées.

3.3 Analyse des données

Tel que mentionné plus haut, l'intervention a été menée dans deux sections du cours FLS 2751. Le corpus analysé ne comporte cependant que les données des étudiants ayant signé le formulaire de consentement, ayant partagé leurs fils de discussion avec ChatGPT et ayant répondu aux deux questionnaires (celui prêtâche et celui post-tâche). Au total, 24 étudiants ont répondu à ces exigences, mais en raison de problèmes techniques, on a perdu accès aux fils de discussion de deux étudiants. Conséquemment, cette étude compte 22 participants. De ces 22 participants, neuf font partie de la collecte de données menée pendant l'intervention en ligne et les treize autres participants font partie de la collecte de données menée pendant l'intervention en mode hybride avec la majorité des participants en présentiel.

Pour analyser les fils de discussion, il fallait d'abord classer les requêtes envoyées par les étudiants et les réponses de ChatGPT. Une analyse taxonomique (Bilmes, 2009) a été utilisée pour classer les messages selon les processus discursifs repérés dans les fils de discussion. L'analyse taxonomique est une approche qui vise la catégorisation sémantique des différentes parties d'une conversation. Bilmes (2009) affirme que pendant une interaction conversationnelle, le sens n'est pas créé par une seule participante, mais il est plutôt coconstruit par les interlocuteurs. Étant donné qu'on s'intéresse en particulier aux interactions entre les apprenants et ChatGPT, l'analyse taxonomique s'est avérée utile pour classer et organiser les requêtes des étudiants et les réponses de la machine.

L'analyse taxonomique menée se situe dans le cadre de l'analyse qualitative du contenu (AQC), ce qui est une approche analytique utilisée pour l'interprétation subjective du contenu de manière qualitative (par le biais de la création des catégories) ainsi que de manière quantitative (en raison de l'analyse de la fréquence des catégories) (Selvi, 2019).

Le logiciel NVivo a été utilisé pour analyser les fils de discussion afin de classer les différents types de requêtes envoyées et les réponses obtenues. Pour ce faire, on a d'abord converti les fils de discussion en documents Word qu'on a ensuite importés dans NVivo. On a décortiqué chaque conversation en tours de parole et l'on a ensuite classé ces tours de parole selon des codes liés aux différents processus discursifs employés, tels que les salutations, le fait de poser une question précise, les demandes de clarification, les négociations de sens, etc. Pour classer les messages, 110 codes ont été employés parmi lesquels 52 ont été utilisés pour classer les requêtes des participants et 58 ont été utilisés pour les réponses de ChatGPT. Puis, ces codes ont été regroupés en catégories thématiquement semblables. Enfin, on a obtenu des données quantitatives en faisant le décompte du nombre de requêtes et de réponses dans chaque catégorie.

Les réponses aux deux questionnaires ont été collectionnées en ligne sur la plateforme Microsoft Forms de manière automatique. Le premier questionnaire ne comprenait que des questions à choix multiples, donc en comptant le nombre de différentes réponses à chaque question, on a obtenu des données quantitatives. Pour le deuxième questionnaire, on a fait de même pour les questions à choix multiples, mais ce deuxième questionnaire comptait aussi deux questions à réponse courte ; la première demandant aux participants de lister les requêtes ayant sollicité des rétroactions utiles et la deuxième demandant aux apprenants de répondre seulement s'ils avaient continué à utiliser ChatGPT pour l'autoapprentissage après l'intervention. L'avant-dernière question était une question à choix multiples qui demandait aux apprenants s'ils avaient continué à utiliser ChatGPT pour l'autoapprentissage après l'expérience. Ce questionnaire était originalement conçu pour être distribué quelques semaines après l'intervention ponctuelle, mais en raison de contraintes de temps, les participants y ont répondu immédiatement après l'intervention. Conséquemment, on ne prend pas en compte les deux dernières questions pour l'analyse des données. Afin d'analyser les

réponses à la question ouverte dans le deuxième questionnaire, on a repris la méthode utilisée pour l'analyse des fils de discussion où l'on a regroupé les réponses thématiquement semblables pour créer des catégories.

Pour ce recueil de données quantitatives et qualitatives, nous adoptons une posture épistémologique qui valorise les deux types de données. Ainsi, nous optons pour une approche post-positiviste, qui encourage la triangulation des méthodes qualitatives et quantitatives, et combine les approches positiviste (visant une analyse objective) et interprétative (visant une analyse subjective) (Panhwar, Ansari & Shah, 2017).

4 Résultats

Cette section présente les résultats de l'étude. D'abord, les résultats du premier questionnaire (le questionnaire préintervention) sont résumés. Ensuite, on présente les résultats obtenus en analysant les fils de discussion avec ChatGPT. Enfin, on détaille les résultats du questionnaire post-intervention. De potentielles implications des résultats sont abordées dans le prochain chapitre.

4.1 Résultats du questionnaire préintervention

Le questionnaire préintervention portait surtout sur les pratiques de révision textuelle des participants avec l'accent mis sur leur usage des outils numériques. Ce questionnaire comportait onze affirmations que les participants ont évaluées en utilisant les échelles de Likert.

La majorité des 22 participants ont déclaré corriger les textes qu'ils écrivent en français soit souvent (59,1 %⁹), soit toujours (31,8 %). Un participant (4,5 %) a déclaré parfois corriger ses textes en français et un participant (4,5 %) a indiqué qu'il ne corrige jamais ses textes en français.

La majorité des participants ont aussi déclaré utiliser des outils numériques plus que des outils traditionnels (tels que les dictionnaires, les Bescherelle en papier, etc.) pendant la révision textuelle.

Un participant (4,5 %) a déclaré toujours se servir des outils traditionnels, quatre participants (18,2 %) ont déclaré souvent s'en servir, sept (31,8 %) ont déclaré parfois s'en servir, sept (31,8 %) ont déclaré rarement s'en servir et trois participants (13,6 %) ont déclaré ne jamais se servir d'outils traditionnels pendant la révision textuelle. En revanche, la majorité des participants ont déclaré utiliser les outils numériques pendant la révision textuelle soit toujours (36,4 %) soit souvent (40,9 %). Deux participants (9,1 %) ont déclaré parfois s'en servir et deux autres participants (9,1 %) ont déclaré ne jamais s'en servir.

⁹ Les pourcentages sont arrondis au dixième près.

Parmi les outils numériques auxquels les participants ont recours pour la révision textuelle, 95,5 % ont déclaré utiliser Microsoft Word, Google Docs ou un autre logiciel de traitement de texte, 72,7 % ont déclaré avoir recours à BonPatron et 18,2 % ont déclaré avoir recours à Antidote. 68,2 % des participants ont aussi signalé qu'ils se servent d'autres outils, mais ces derniers n'étaient pas identifiés.

La cinquième affirmation dans le questionnaire déclarait « je suis satisfait des outils de correction linguistique numériques que j'utilise ». 13,6 % des participants étaient d'accord avec cette affirmation, 72,7 % ont déclaré être plutôt d'accord avec cette affirmation et 13,6 % ont déclaré des sentiments neutres. Aucun participant n'a déclaré être « plutôt pas d'accord » ou « pas d'accord » et aucun participant n'a choisi la réponse « ne s'applique pas, je n'utilise pas d'outils de correction linguistique numériques. » La prochaine affirmation était « j'ai confiance dans les corrections des outils de correction linguistique numériques que j'utilise ». 36,4 % des participants ont été d'accord, 50 % étaient plutôt d'accord, 9,1 % étaient neutres et 4,5 % avaient choisi l'option « plutôt pas d'accord. »

La septième affirmation du questionnaire était « je comprends les rétroactions des outils de correction linguistiques numériques que j'utilise ». 13,6 % des participants étaient d'accord avec cet énoncé, 50 % étaient plutôt d'accord, 9,1 % étaient neutres et 13,6 % des participants ont déclaré être plutôt pas d'accord. La prochaine affirmation visait à évaluer l'intérêt des participants à la possibilité d'interagir avec les outils de correction numériques dont ils se disposent. La majorité des participants était d'accord (27,3 %) ou plutôt d'accord (36,4 %) de vouloir interagir avec ces outils. 18,2 % des participants étaient neutres, 9,1 % étaient plutôt pas d'accord et 9,1 % n'étaient pas d'accord.

Les deux prochaines affirmations portaient sur l'utilisation des outils d'intelligence artificielle pendant la rédaction et la révision textuelle des textes français. La moitié des participants ont déclaré utiliser rarement (9,1 %) ou ne jamais (40,9 %) utiliser les outils d'IA pour écrire en français. L'autre moitié des participants ont déclaré parfois (27,3 %) ou souvent (22,7 %) utiliser ces outils, mais aucun participant n'a déclaré toujours utiliser les outils d'IA pour la rédaction en français. Pour l'affirmation portant sur l'utilisation des outils d'IA pour la correction des textes français ; un peu plus de la moitié des étudiants ont déclaré rarement (13,6 %) ou ne jamais (40,9 %) se servir d'outils d'IA pour la correction. 13,6 % des participants ont déclaré parfois utiliser ces outils, 22,7 % ont déclaré souvent s'en servir et seulement 9,1 % des participants ont déclaré toujours utiliser les outils d'IA pour corriger leurs textes rédigés en français.

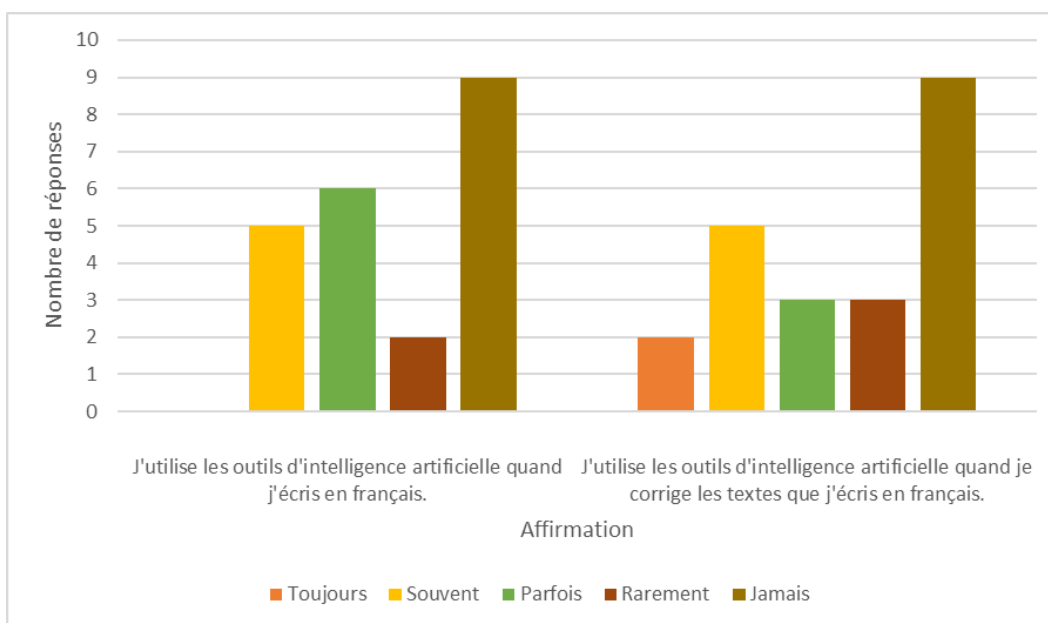


Figure 9 - Les réponses aux questions portant sur l'utilisation de l'IA pour la rédaction et la révision textuelle en français

La dernière affirmation avait pour but de jauger les attitudes des participants envers l'IA. La moitié des participants (50 %) ont déclaré avoir des opinions plutôt positives de l'IA, un participant (4,5 %) a déclaré avoir une opinion positive et cinq participants (22,7 %) avaient des sentiments

neutres. Trois participants (13,6 %) avaient des opinions plutôt négatives et deux participants (9,1 %) avaient des opinions négatives envers l'IA. Ce qui est intéressant de noter est que tous les participants ayant déclaré des opinions négatives et plutôt négatives envers l'IA ont indiqué ne jamais s'en servir pour écrire ou pour corriger leurs textes français, exception faite d'un participant avec une opinion plutôt négative envers l'IA qui a déclaré s'en servir souvent pour écrire en français.

En somme, on constate que la majorité des participants corrigent les textes qu'ils écrivent en français et pour ce faire, ils ont surtout recours aux outils de correction linguistique numériques. Les participants ont déclaré avoir des perceptions généralement positives de ces outils. Enfin, la moitié des participants ne se servent pas (ou se servent peu) des outils d'IA pour écrire en français et la majorité des participants ne s'en servent pas pour la correction linguistique. Bien que l'utilisation des outils d'IA pour la rédaction et pour la révision textuelle en français ne soit pas répandue parmi les participants, la majorité (54,5 %) déclare des opinions favorables envers les outils d'IA.

4.2 Résultats de l'analyse des fils de discussion d'interaction avec ChatGPT

L'analyse des fils de discussion comporte plusieurs étapes ; d'abord une analyse pour obtenir des données quantitatives portant sur les fils de discussion, comme le nombre de mots par fils de discussion, le nombre de tours de parole, etc. On a ensuite mené une analyse des messages produits par les étudiants (les requêtes) ainsi qu'une analyse des messages envoyés par ChatGPT (les réponses). On a analysé l'alignement requête-réponse pour voir quels types de requêtes ont sollicité quels types de réponses. Enfin, on s'est intéressé aux réponses de ChatGPT qui étaient incorrectes ou inappropriées en contexte.

4.2.1 Analyse quantitative des fils de discussion

Le Tableau 4 présente le nombre de mots total dans les fils de discussion, le nombre de tours de parole, le nombre de requêtes imposées utilisées, le nombre de requêtes additionnelles formulées par les participants, ainsi que les quantités moyennes et les quantités totales. Le modèle conversationnel que suit ChatGPT dicte qu'il ne peut pas entamer une conversation ; c'est l'utilisateur qui doit envoyer le premier message. ChatGPT répond à chaque message de l'utilisateur avec une réponse dans un seul message. La machine est donc toujours la dernière à envoyer un message dans un fil de discussion. Pour cette raison, les tours de paroles sont toujours en nombres pairs.

Rappelons par ailleurs que deux groupes d'étudiants ont participé à la collecte de données. P1-P9 ont participé à l'intervention en ligne et P10-P22 ont assisté à l'intervention hybride avec la grande majorité des étudiants en présentiel. Les deux groupes avaient environ une heure pour réviser leur texte avec ChatGPT. Les résultats montrent que les participants à la collecte de données menée pendant la deuxième intervention, celle en classe (désormais Groupe 2) se sont impliqués plus que ceux ayant participé à la première intervention, celle en ligne (désormais Groupe 1). Dans le Tableau 4, les données des participants du Groupe 1 sont surlignées en violet pour permettre une distinction plus évidente entre les deux groupes.

Participant	Nombre total de mots	Tours de parole	Requêtes imposées	Nouvelles requêtes
P1	494	6	2	1
P2	894	12	2	4
P3	399	6	2	1
P4	402	6	2	1
P5	452	4	2	0
P6	670	8	2	2
P7	579	10	1	3
P8	578	8	2	2

P9	544	4	2	0
P10	544	10	2	3
P11	1260	14	2	5
P12	990	10	2	3
P13	1205	12	2	4
P14	468	4	2	0
P15	485	6	2	1
P16	956	12	2	4
P17	660	10	1	4
P18	1656	18	2	7
P19	1370	20	2	8
P20	1292	14	2	5
P21	645	10	2	2
P22	788	16	1	7
Total	17331	220	41	67
Moyenne	787,8	10	1,9	3

Tableau 4 - Tableau de synthèse des données quantitatives des fils de discussion

Le corpus analysé comprend au total 17331 mots qui constituent 220 tours de paroles (110 requêtes des étudiants et 110 réponses de ChatGPT). Rappelons que seulement deux requêtes ont été imposées ; « révise ce texte » et « explique toutes les révisions ». Tous les participants ont employé ces requêtes sous une forme ou une autre, sauf P7 qui a omis la deuxième requête imposée demandant à ChatGPT d'expliquer toutes ses modifications et P17 et P22 qui n'ont pas demandé une explication des révisions, mais une liste des modifications. Deux participants (P7 et P21) ont envoyé la requête « révise ce texte » sans inclure le texte à réviser, ce qui a ajouté deux tours de parole à leurs fils de discussion. Toutefois, la requête « révise ce texte » et le texte envoyé seul sont considérés des requêtes imposées et ne figurent donc pas dans les nouvelles requêtes. Le fait d'avoir divisé une des requêtes imposées en deux messages n'est pas l'équivalent d'envoyer deux requêtes imposées. Conséquemment, on compte 41 occurrences de l'emploi des requêtes imposées, mais 43 messages envoyés pour ce faire. Pour la colonne des requêtes imposées dans le Tableau

4, on compte le nombre d'occurrences d'emploi des requêtes imposées et non le nombre de messages utilisés.

Le plus long fil de discussion était celui de P18, comptant 1656 mots. Le plus court était celui de P3 avec 399 mots. En moyenne, les fils de discussion comptaient 787,7 mots. Le fil de discussion de P19 comptait le plus grand nombre de tours de parole (20), comparé aux fils de discussion de P5, P9 et P14 qui ne comptaient que quatre tours de parole chacun. Ces trois participants n'ont employé que les deux requêtes imposées, n'inventant aucune nouvelle requête. Conséquemment, les fils de discussion de ces participants ne comptent que quatre tours de paroles chacun ; deux requêtes imposées et deux réponses de ChatGPT. En moyenne, les fils de discussion comportent 10 tours de parole.

Tel que dit en haut, tous sauf trois participants ont employé les deux requêtes imposées. Conséquemment, en moyenne, les participants ont employé 1,9 requête imposée. Les participants ont créé un total de 67 nouvelles requêtes, c'est-à-dire des requêtes additionnelles qui n'ont pas été imposées. En moyenne, les fils de discussion comprennent trois requêtes additionnelles créées par les participants. Des requêtes additionnelles, celles envoyées les plus souvent étaient différentes versions de la requête « je ne comprends pas [révision] », visant à solliciter de manière implicite une explication additionnelle d'une révision proposée par ChatGPT. Cinq participants ont envoyé cette requête pour un total de huit occurrences ; un participant l'avait envoyée deux fois et un autre l'avait envoyée trois fois. La majorité des requêtes additionnelles étaient des requêtes discrètes et différentes, mais la prochaine sous-section les regroupe en catégories selon l'objectif du message (obtenir une explication additionnelle, remercier la machine, etc.) pour offrir une vue d'ensemble des types de requêtes formulées.

4.2.2 Analyse des requêtes formulées par les apprenants

Le Groupe 1 a créé quatorze nouvelles requêtes, dont huit portaient sur la révision textuelle. Les six autres requêtes étaient des salutations, des remerciements et des questions hors sujet. Le Groupe 2 a créé 53 nouvelles requêtes. Parmi ces requêtes additionnelles, 45 portaient sur la révision textuelle et les 8 autres nouvelles requêtes étaient des messages autres tels que des salutations et des remerciements.

Groupe	Nouvelles requêtes liées à la révision textuelle	Requêtes autres	Total	Pourcentage de requêtes liées à la révision textuelle arrondi au dixième près
Groupe 1	8	6	14	57,1 %
Groupe 2	45	8	53	84,9 %
Au total	53	14	67	79,1 %

Tableau 5 - Tableau de synthèse des nouvelles requêtes créées par les deux groupes de participants

Il y a un écart considérable entre les efforts d'engagement des deux groupes. Cet écart pourrait être expliqué par le fait que le Groupe 1 a participé à l'intervention en ligne, tandis que la majorité des participants dans le Groupe 2 étaient en présentiel avec la chercheuse (l'auteure). Il semble donc que les étudiants ayant participé à l'intervention en personne se sont plus impliqués pendant la révision textuelle avec ChatGPT. En moyenne, les fils de discussion du Groupe 1 comptent 556,9 mots et 7,1 tours de parole, tandis que ceux du Groupe 2 comptent 947,6 mots et 12,2 tours de parole. D'ailleurs, 84,9 % des requêtes additionnelles formulées par le Groupe 2 étaient liées à la révision textuelle, contre 57,1 % pour le Groupe 1.

Il est aussi intéressant de noter que quatre participants (P1, P3 et P4 du Groupe 1 et P15 du Groupe 2) n'ont formulé qu'une requête additionnelle. Les requêtes additionnelles de P1 et de P3 portaient sur la révision, tandis que les nouvelles requêtes de P4 et de P15 étaient des requêtes de socialisation (un remerciement et une salutation, respectivement). P1 a posé une question portant

sur l'importance du questionnement direct après que ChatGPT a modifié son texte pour l'employer et P3 a sollicité une précision sur une des révisions de ChatGPT. En revanche, P4 a remercié la machine après avoir obtenu sa révision et P15 a entamé son fil de discussion en saluant la machine.

Bien que certains participants aient été plus engagés que d'autres pendant la révision textuelle avec ChatGPT, il y a toujours des tendances présentes et on a tenté de les relever en classant les tours de parole dans les fils de discussion selon des codes représentant les différents processus discursifs employés, tels que les salutations, le fait de poser une question précise, les demandes de clarification, les négociations de sens, etc. L'étape du codage initial a permis de relever les processus discursifs employés dans les messages et d'associer à chaque différent message une brève description, ce qui a facilité la prochaine étape ; la création de catégories. Les catégories ont été créées afin de donner une vue d'ensemble des différents types de requêtes formulées et des différentes réponses obtenues. La typologie utilisée pour catégoriser les requêtes des étudiants a été élaborée sur la base des objectifs que les participants tentaient d'accomplir en envoyant chaque message. Le Tableau 6 présente les différentes catégories utilisées pour classer les requêtes des étudiants, une courte description de chacune, le nombre de requêtes classées dans chaque catégorie, ainsi que les pourcentages.

Catégorie	Description	Nombre de requêtes	Pourcentage
Requête imposée	Les deux requêtes imposées : (révise ce texte et explique toutes les révisions).	43	39,1 %
Clarification	Les requêtes sollicitant des clarifications et des explications additionnelles sur les éléments dans le texte révisé, dans le texte original, ou concernant les deux à la fois.	42	38,2 %
Socialisation avec la machine	Les requêtes saluant, remerciant ou complimentant ChatGPT.	11	10 %
Désaccord	Les requêtes où le participant exprime son désaccord avec ChatGPT.	6	5,5 %

Question autre	Les requêtes posant des questions non liées à la révision textuelle.	3	2,7 %
Révision	Les requêtes sollicitant une révision additionnelle ou portant sur une révision additionnelle.	3	2,7 %
Modification	Les requêtes demandant à ChatGPT de modifier le texte.	2	1,8 %

Tableau 6 - Typologie pour classer les requêtes étudiantes

La catégorie avec le type de requête le plus fréquemment employé par l'ensemble des participants est celle des « Requêtes imposées » (43 occurrences). Cette catégorie renferme les deux requêtes imposées sollicitant une révision initiale ainsi qu'une explication de cette dernière. Ensuite, il y a la catégorie « Clarification » qui comporte 42 requêtes. Des exemples de requêtes dans cette catégorie sont : « Je ne comprend pas correction #5 [*sic*] » visant à obtenir une explication additionnelle portant sur le texte révisé. Une autre requête dans cette catégorie est « peux-tu explique la difference entre connaissiez et sauriez? [*sic*] » où l'apprenant sollicite une clarification après que ChatGPT a remplacé le verbe « connaissiez » par le verbe « sauriez » pendant la révision.

La catégorie « Socialisation avec la machine » compte onze requêtes, telles que les salutations comme « bonjour » et les remerciements comme « Merci (: ». Les six requêtes exprimant un désaccord de la part de l'étudiant sont classées dans la catégorie « Désaccord ». Dans cette catégorie, il y a par exemple, la requête « mais je pense que la verbe écouter est mieux que entends dans cette situation [*sic*] » qui exprime un désaccord de la part de l'étudiant qui essaie d'opposer une révision faite par ChatGPT qui avait remplacé le verbe « écouter » par le verbe « entendre ».

La catégorie « Question autre » regroupe trois requêtes où le participant pose une question qui n'est pas reliée à la révision textuelle. Par exemple, un participant a envoyé la requête « Explique le mot “Slay” », sollicitant une définition d'un mot anglais. Les trois requêtes portant sur des révisions autres que celle initiale sont regroupées dans la catégorie « Révision ». La requête « Est-

ce que tu peux réviser le texte maintenant? *[sic]* » est un exemple d'une requête dans cette catégorie où le participant sollicite une révision additionnelle.

Enfin, la catégorie « Modification » classe deux requêtes où un participant a dit à ChatGPT de modifier son texte. Un participant a envoyé la requête « Remplace “salut à tous” avec 10 », essayant d'inciter ChatGPT à modifier son texte en remplaçant sa phrase d'introduction par le dixième élément dans une liste d'expressions alternatives.

4.2.3 Analyse des réponses fournies par ChatGPT

Neuf catégories ont été créées pour classer les différentes réponses de ChatGPT selon le contenu du message. Ces catégories, leurs descriptions, le nombre de réponses qu'elles regroupent et les pourcentages sont présentés dans le Tableau 7.

Catégorie	Description	Nombre de Réponses	Pourcentage
Réponse attendue	Les réponses de ChatGPT répondant aux deux requêtes imposées.	43	39,1 %
Clarification	Les réponses fournissant une explication ou une clarification additionnelle sur un élément dans le texte révisé, dans le texte original, ou concernant les deux à la fois.	33	30 %
Socialisation avec le participant	Les réponses exprimant les salutations, les remerciements, etc.	11	10 %
Liste	Les réponses fournissant une liste d'éléments.	7	6,4 %
Concession	Les réponses où ChatGPT concède après qu'un participant exprime un désaccord ou pose une question.	4	3,6 %
Modification	Le texte avec des modifications autres que celles effectuées pendant la révision initiale.	4	3,6 %
Autocorrection	Les réponses où ChatGPT se corrige.	4	3,6 %
Réponse autre	Les réponses répondant aux questions hors sujet.	3	2,7 %
Révision	Les réponses fournissant le texte révisé une deuxième fois.	1	1 %

Tableau 7 - Typologie pour classer les réponses de ChatGPT

La catégorie qui regroupe le plus grand nombre de réponses de ChatGPT est « Réponse attendue ». Cette catégorie contient les 43 réponses aux requêtes imposées telles que le texte révisé (Figure 10) et une liste des révisions (Figure 11). Afin de mieux contextualiser les réponses, les figures dans cette section de l'analyse comportent les réponses de ChatGPT ainsi que les requêtes des étudiants les ayant sollicitées. D'ailleurs, les figures comportant des réponses avec des fautes sont marquées d'un astérisque. Les réponses incorrectes et inappropriées sont traitées dans la prochaine sous-section.

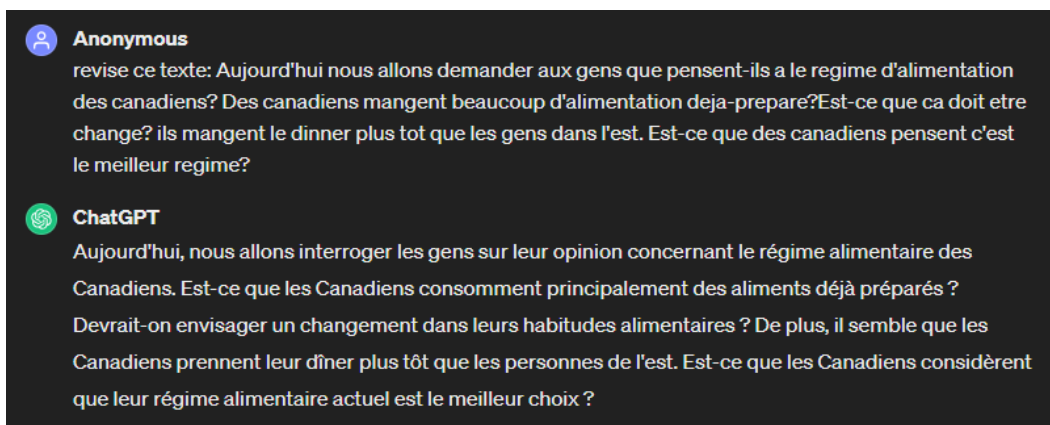


Figure 10 - Une réponse dans la catégorie « Réponse attendue » à la première requête imposée

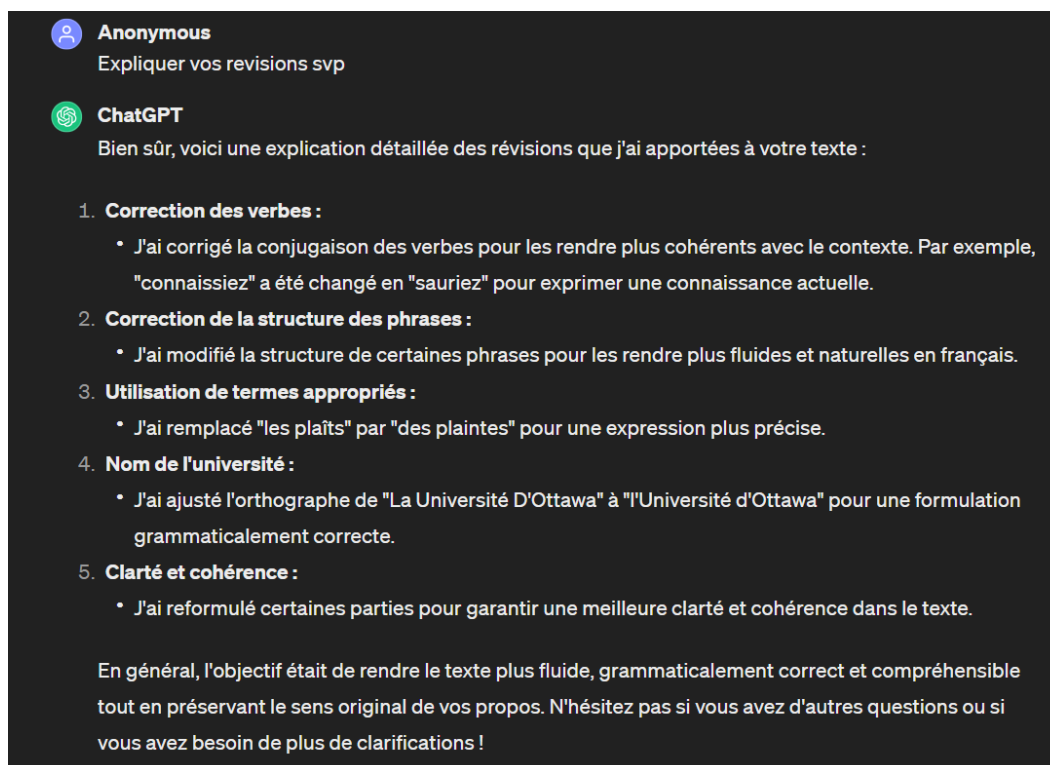




Figure 11 - Une réponse dans la catégorie « Réponse attendue » à la deuxième requête imposée

Il y a ensuite la catégorie « Clarification » qui classe 33 réponses qui apportent des explications et des clarifications additionnelles sur les éléments dans le texte révisé (Figure 12), les éléments dans le texte original (Figure 13), ou concernant les deux textes à la fois (Figure 14).


 **Anonymous**
ok. mais pourquoi as-tu changer écoutais avec entends?


 **ChatGPT**
Je me suis permis de changer "j'écoutais" en "j'entends" pour deux raisons principales :

1. **Présent de l'indicatif pour l'actualité** : En général, le présent de l'indicatif est utilisé pour exprimer des faits actuels et des vérités générales. En disant "j'entends", cela donne une dimension d'actualité à votre déclaration. Vous mentionnez que vous avez souvent entendu des plaintes, et le présent suggère que ces plaintes sont quelque chose que vous continuez à percevoir actuellement.
2. **Cohérence du temps** : En ajustant la phrase, j'ai cherché à maintenir la cohérence du temps dans le texte. Vous parlez de quelque chose qui semble être une observation continue, donc l'utilisation du présent est plus appropriée pour rendre compte de cette observation régulière.

Cependant, si vous préférez conserver le passé, vous pourriez dire "j'écoutais souvent" pour indiquer que c'était une habitude ou une activité fréquente dans le passé. La modification que j'ai apportée était principalement pour souligner la continuité et l'actualité de l'écoute des plaintes.

Figure 12 - Une réponse dans la catégorie « Clarification » liée à une requête de clarification

 **Anonymous**
est-ce que l'utilisation d'imparfait est correct dans ma texte?

 **ChatGPT**
L'utilisation de l'imparfait dans votre texte initial n'est pas tout à fait appropriée. Vous avez utilisé l'imparfait dans la phrase "Si vous habitiez au Ottawa, vous connaissiez qu'il fait si froid en hiver."
L'imparfait est généralement utilisé pour décrire une action habituelle, continue ou en cours dans le passé.

Dans votre contexte, puisque vous faites référence à une connaissance potentielle ou hypothétique qui pourrait exister dans le futur (si la personne habitait à Ottawa), le conditionnel présent serait plus approprié. Ainsi, la phrase pourrait être reformulée comme suit :

"Si vous habitiez à Ottawa, vous sauriez qu'il fait très froid en hiver."

Cette correction rend le verbe "sauriez" cohérent avec le contexte de connaissance potentielle liée à la situation hypothétique d'habiter à Ottawa.

Figure 13 - Une réponse dans la catégorie « Clarification » liée à une requête de clarification

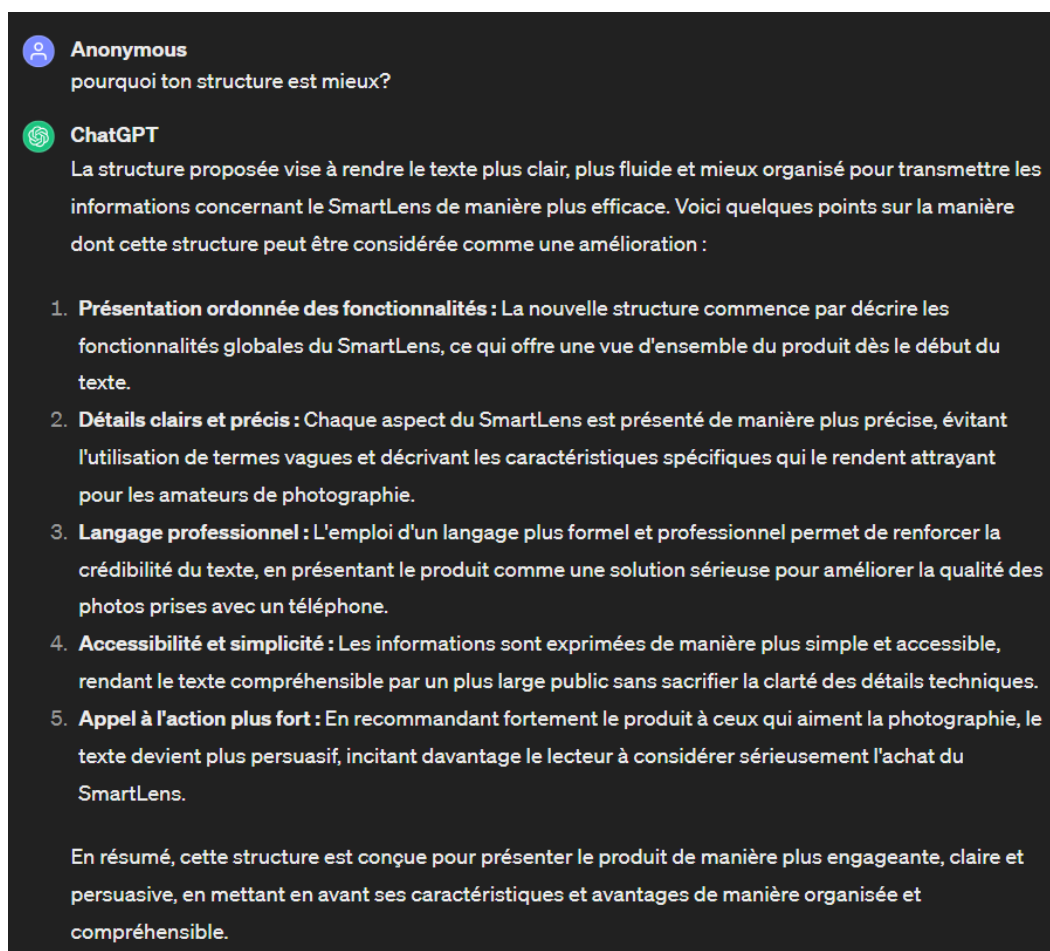


Figure 14 - Une réponse dans la catégorie « Clarification » liée à une requête de clarification

La catégorie « Socialisation avec le participant » comprend onze réponses incluant des salutations et des remerciements (Figure 15).

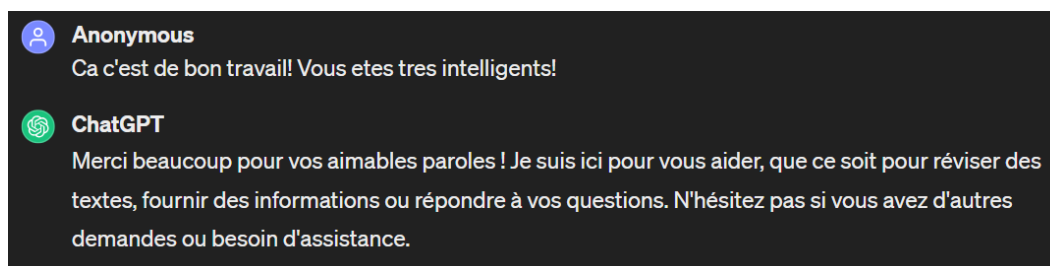


Figure 15 - Une réponse dans la catégorie « Socialisation avec le participant »

La catégorie « Liste » classe les sept réponses où ChatGPT propose une liste d'éléments à l'utilisateur. La Figure 16 est un exemple d'une liste de salutations alternatives proposées par la

machine. Les quatre réponses classées dans la catégorie « Modification » sont les réponses proposant une version modifiée du texte autre que celle envoyée pendant la révision initiale (Figure 17).

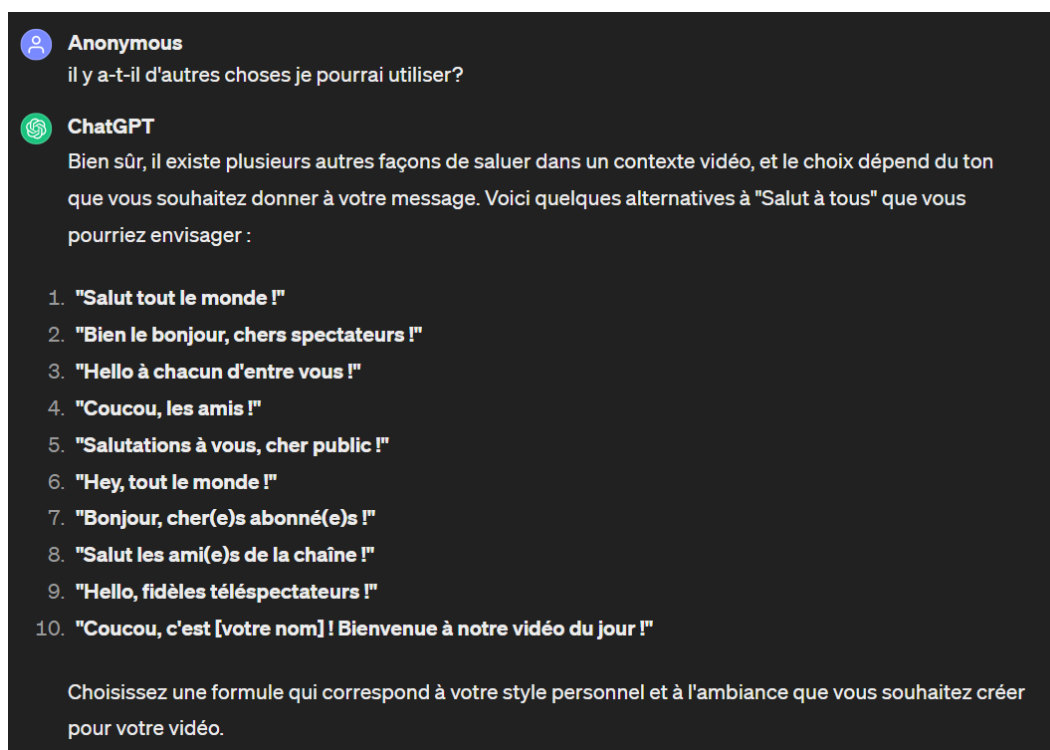


Figure 16 - Une réponse dans la catégorie « Liste » qui répond à une requête sollicitant une information supplémentaire

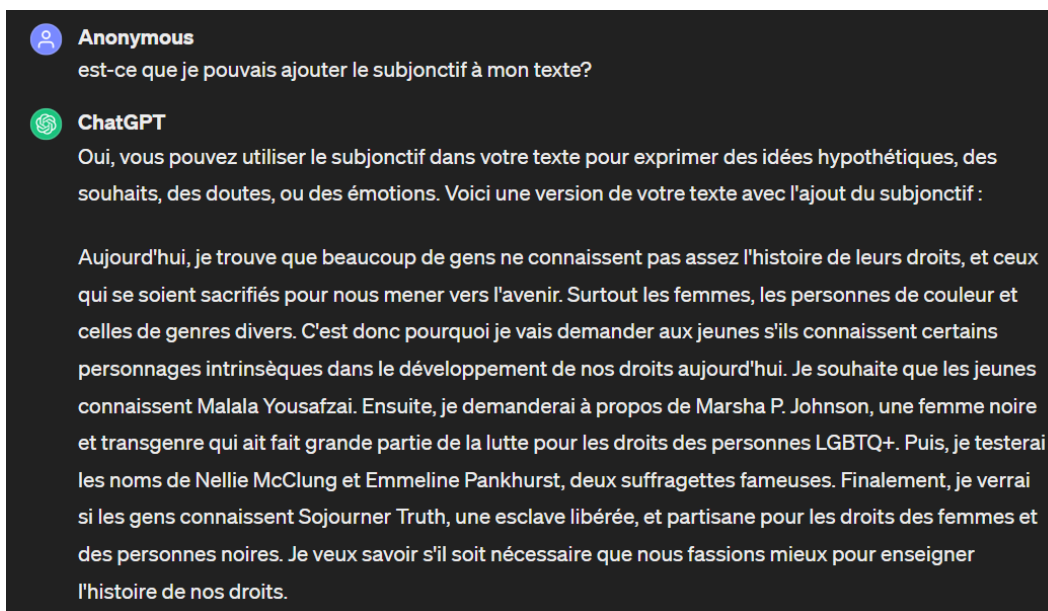


Figure 17* - Une réponse dans la catégorie « Modification » liée à une requête de clarification

ChatGPT s'est autocorrigé quatre fois et ces réponses sont classées dans la catégorie « Autocorrection » (Figure 18). La catégorie « Concession » classe quatre réponses où ChatGPT propose soit une concession, soit un compromis, tel qu'illustré dans la Figure 19. Les trois réponses dans « Réponse autre » sont des réponses aux questions hors sujet posées par les participants (Figure 20). La dernière catégorie, « Révision », classe une seule réponse qui repartage le texte révisé une deuxième fois (Figure 21).

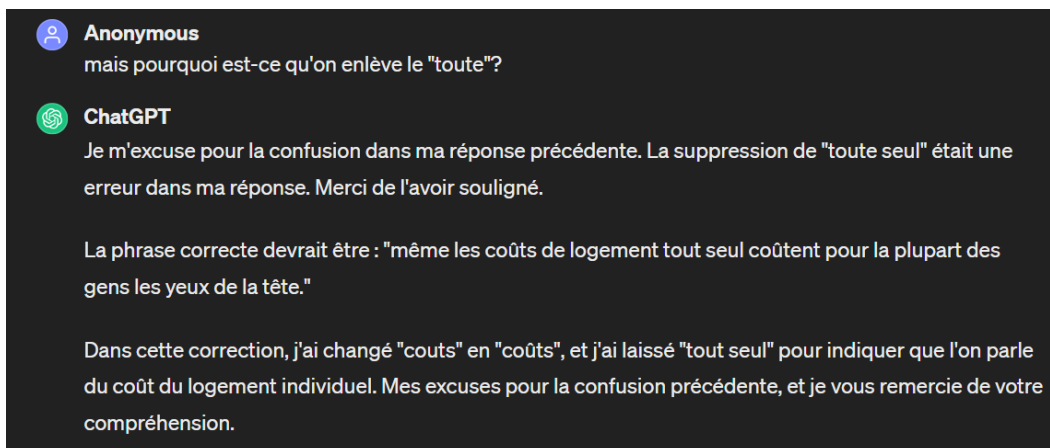


Figure 18 - Une réponse dans la catégorie « Autocorrection » liée à une requête de clarification

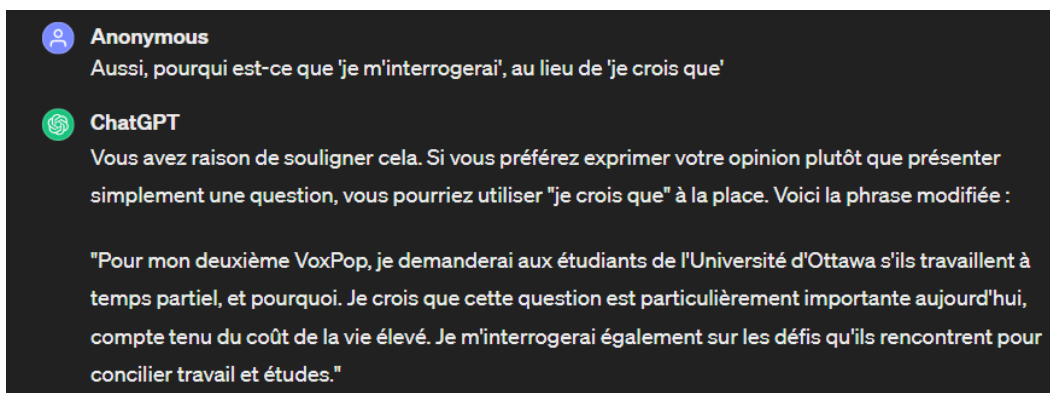


Figure 19 - Une réponse dans la catégorie « Concession » liée à une requête de clarification

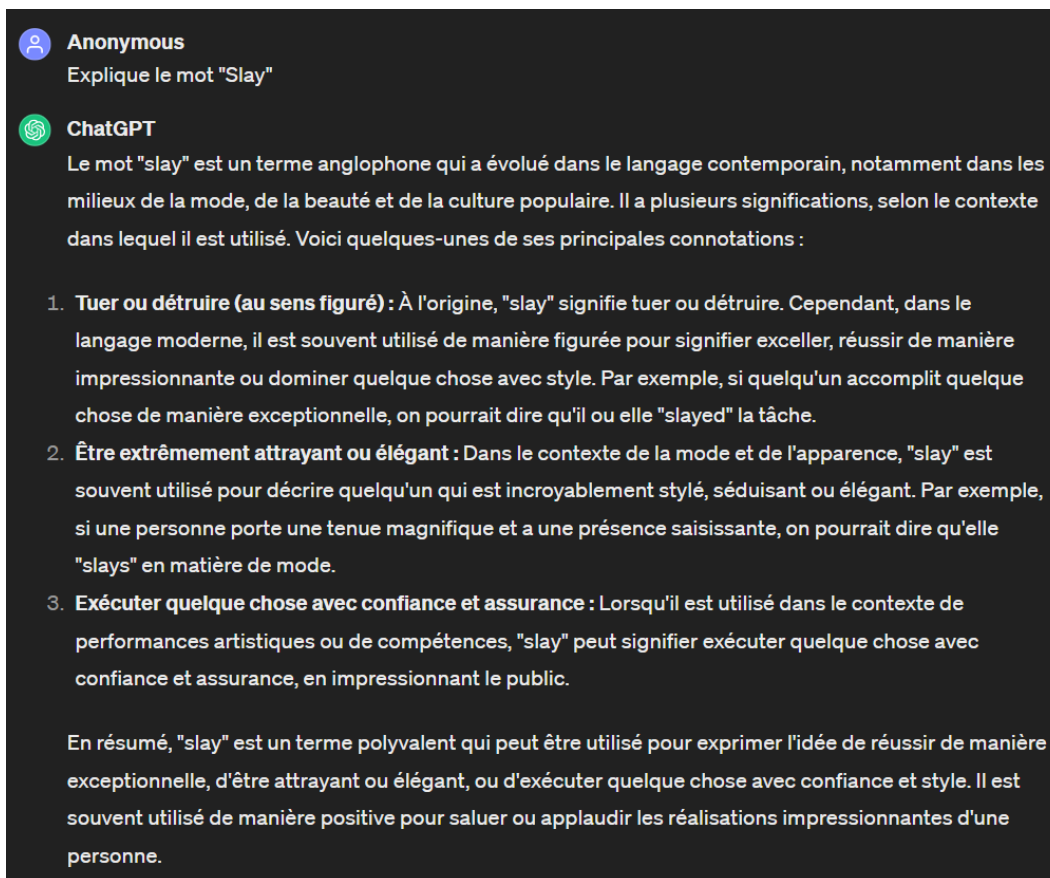


Figure 20 - Une réponse dans la catégorie « Réponse autre » qui répond à une question autre (la définition d'un terme anglais)

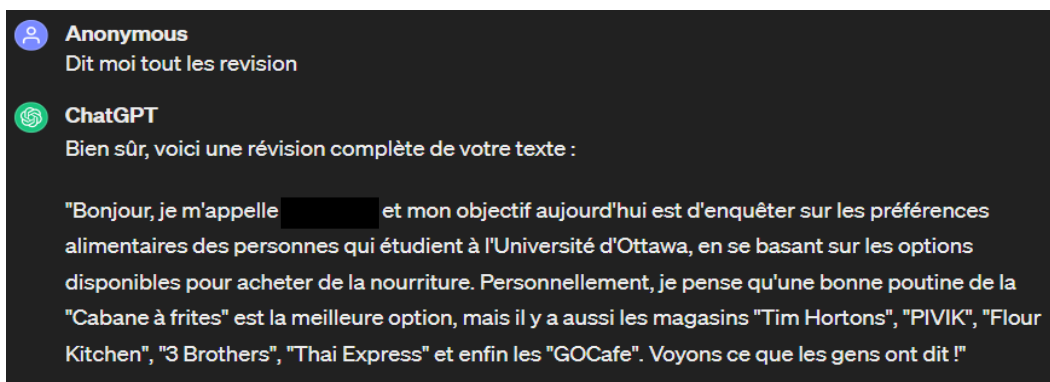


Figure 21* - Une réponse dans la catégorie « Révision » liée à une requête de clarification

4.2.3.1 Analyse de l'alignement requête-réponse

On a aussi étudié quelles catégories de requêtes avaient sollicité quelles catégories de réponses. Il y a des corrélations directes et étroites, telles que la relation entre la catégorie de requêtes « Requête

imposée » qui classe 43 requêtes et la catégorie de réponses « Réponses attendues » qui classe 43 réponses. Nécessairement si un étudiant a soumis une requête imposée, la réponse a été une réponse attendue. Cette tendance s'est reproduite pour les relations « Socialisation avec la machine-Socialisation avec un participant » où chaque requête de socialisation d'un étudiant a sollicité une réponse de socialisation de la machine. Les catégories « Question autre-Réponse autre » s'alignent aussi. Si l'on enlève les requêtes dans ces catégories, il reste encore 52 requêtes dans les catégories « Clarification », « Désaccord », « Révision », et « Modification » ayant sollicité 52 réponses classées dans les catégories « Clarification », « Liste », « Modification », « Autocorrection », « Concession » et « Révision ».

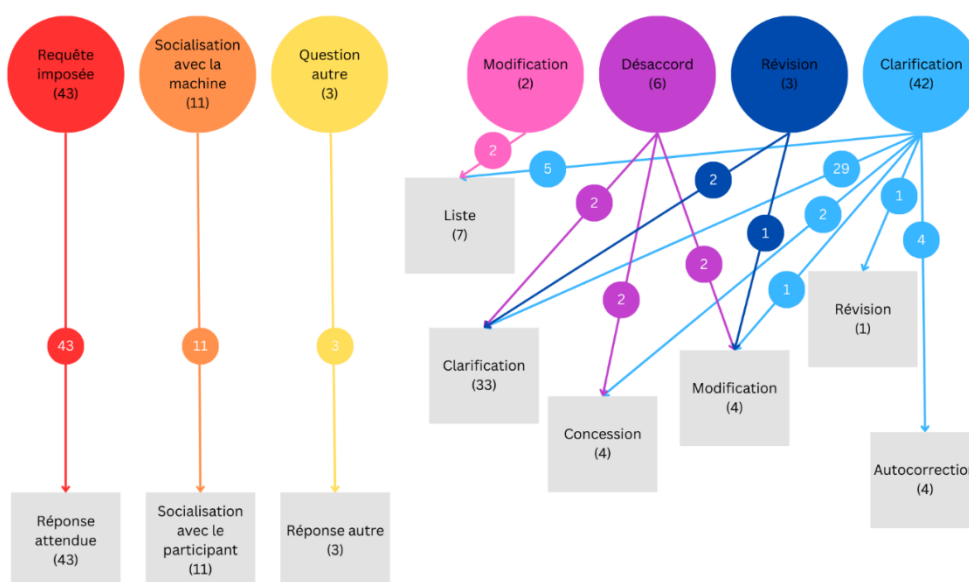


Figure 22 - Diagramme de l'alignement requête-réponse

La Figure 22 est une représentation visuelle de l'alignement requête-réponse qui illustre que parfois les requêtes dans une même catégorie ont sollicité des réponses dans différentes catégories. Par exemple, quand un participant a envoyé une requête de type « Désaccord », ChatGPT a parfois répondu avec une clarification (deux occurrences), une concession (deux occurrences) ou une

modification (deux occurrences). Il semble donc que ChatGPT est en mesure de répondre aux requêtes en adoptant une démarche de cas par cas. Par exemple, remettre en question ce que propose la machine n'engendre pas toujours une concession, ce qui suggère que ChatGPT peut prendre en considération l'information présente dans une requête et s'en servir en contexte afin de proposer une réponse logique au lieu de répondre avec un automatisme qui ne serait pas toujours pertinent. Cependant, on constate que les requêtes imprécises ou mal formulées peuvent engendrer des réponses inappropriées ou incorrectes. La prochaine sous-section aborde les réponses inappropriées et incorrectes et propose également des exemples.

4.2.3.2 Analyse des réponses incorrectes et inappropriées fournies par ChatGPT

Bien que les interactions lors de la révision textuelle soient le point de mire de cette étude, une analyse d'ensemble a été menée pour évaluer la qualité des réponses de ChatGPT. Pour analyser si les requêtes étaient appropriées et correctes, on a repris les fils de discussion afin de voir le contexte dans lequel les réponses étaient proposées. De cette manière, on a pu déterminer si une réponse répondait de manière appropriée et pertinente à la requête l'ayant sollicitée. On a aussi pu déterminer si une réponse proposée était correcte selon le contexte.

Cette étape de l'analyse a relevé que des 110 réponses de ChatGPT, 20 (18,2 %) étaient incorrectes ou inappropriées. On s'intéresse surtout aux manières dont les réponses étaient incorrectes ou inappropriées, ainsi que les contextes dans lequel ces réponses apparaissent. Des 20 réponses identifiées, la moitié sont des cas où ChatGPT prétend avoir effectué une modification qui n'a pas été faite ou classe une modification dans une mauvaise catégorie. Ce type de réponse incorrecte n'est pas causée par une requête imprécise ou mal formulée et il s'agit plutôt d'une occasion où la machine s'est trompée. Dans la Figure 23, ChatGPT explique que le mot « bienvenue » a été corrigé pour être accordé avec « à tous ». Il est vrai que ChatGPT a corrigé l'orthographe de ce

mot (dans le texte original, ce mot était épelé « bienvenu », ce qui était incorrect dans le contexte dans lequel il était employé), mais cette modification était apportée pour rectifier l'orthographe et non pour accorder le mot. Dans la Figure 24, ChatGPT fabule (Cf. Bang et collab., 2023 ; Kaplan-Rakowski et collab., 2023) et prétend avoir remplacé « cours de labo » par « cours de laboratoire », bien que le texte original dût déjà « cours de laboratoire ».

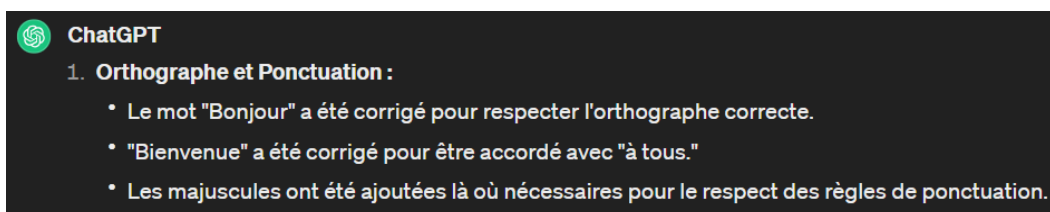


Figure 23 - Extrait d'une réponse incorrecte où ChatGPT ne décrit pas correctement une modification apportée

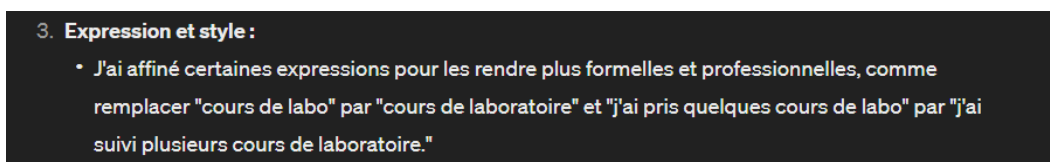


Figure 24 - Extrait d'une réponse incorrecte où ChatGPT prétend avoir effectué une modification qui n'a pas été faite

On dénote deux occasions où ChatGPT a repartagé le texte d'un étudiant en conservant un élément précis, mais a aussi répété des fautes précédemment révisées. La Figure 25 est un exemple où le participant a voulu conserver le verbe « écouter ». ChatGPT explique qu'il est possible d'employer ce verbe et partage le texte en l'employant, mais inclut des erreurs et des phrases maladroites qui avaient été rectifiées pendant la révision initiale. Deux exemples sont soulignés dans la Figure 25.

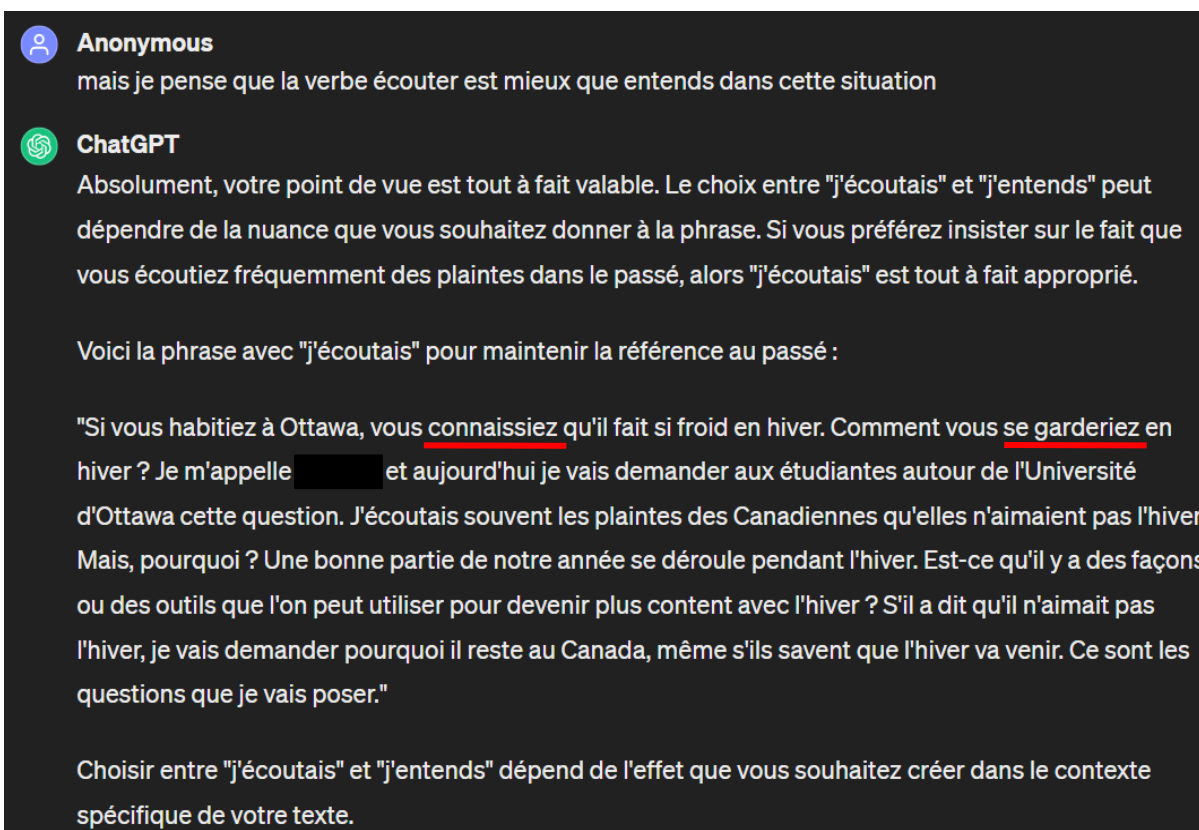


Figure 25 - Une réponse incorrecte où ChatGPT repartage le texte avec des erreurs précédemment corrigées

Il y a aussi une instance où ChatGPT a commis une erreur en modifiant le texte. Dans la Figure 26, le participant avait posé une question pour savoir s'il était possible d'employer le subjonctif dans son texte. ChatGPT a répondu avec le texte réécrit pour inclure des verbes au subjonctif. Cependant, l'emploi du subjonctif dans le texte modifié n'est pas toujours correct. L'extrait souligné en rouge est un exemple où l'emploi du subjonctif est incorrect.

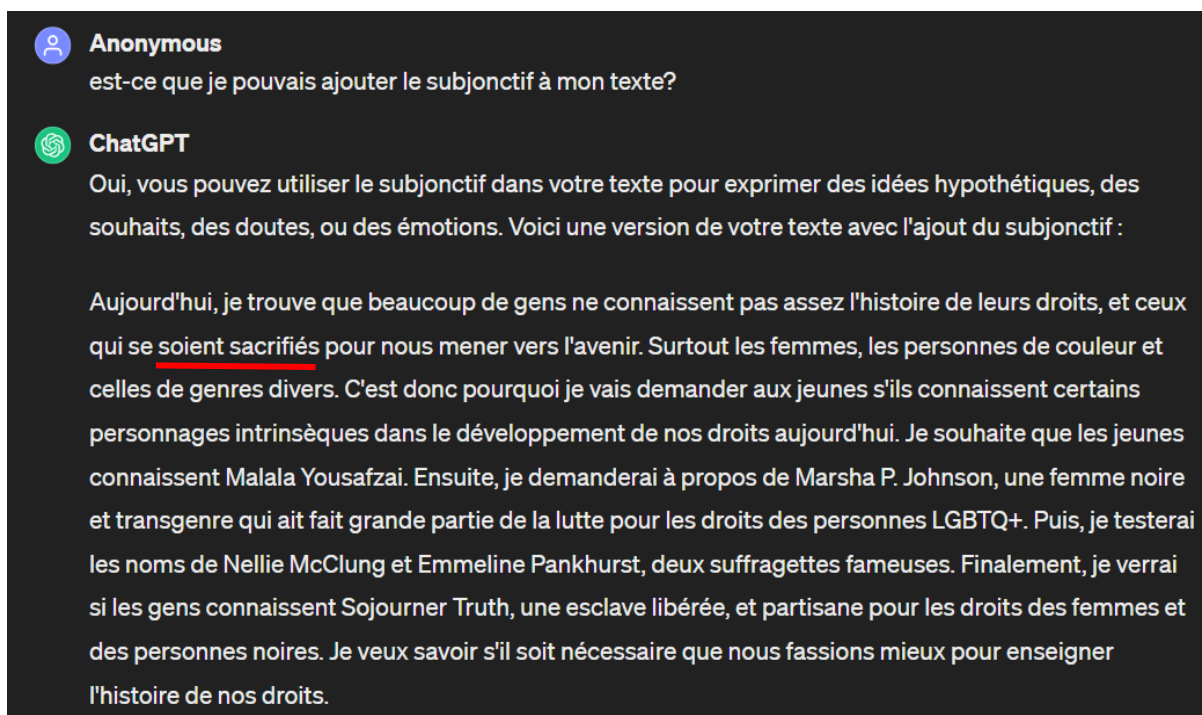


Figure 26 - Une réponse incorrecte où ChatGPT commet une erreur de conjugaison

Certaines réponses ont été identifiées comme étant inappropriées en raison d'une incompréhension de la part de la machine. Ces incompréhensions sont parfois causées par des requêtes imprécises ou mal formulées et sont parfois en raison d'une simple incompréhension de ChatGPT. La Figure 27 s'agit d'un cas où ChatGPT semble avoir mal interprété une requête d'un participant. Dans cet exemple, le participant demande à ChatGPT d'effectuer une modification dans son texte en remplaçant sa phrase d'introduction « Salut à tous » par le dixième élément dans une liste d'expressions alternatives proposées par ChatGPT dans la réponse précédente. La requête formulée est trop succincte, ce qui engendre une mécompréhension de la part de ChatGPT. Au lieu de modifier le texte en remplaçant « Salut à tous » par l'élément désiré, ChatGPT répond avec la même liste d'expressions alternatives, mais réordonnées.

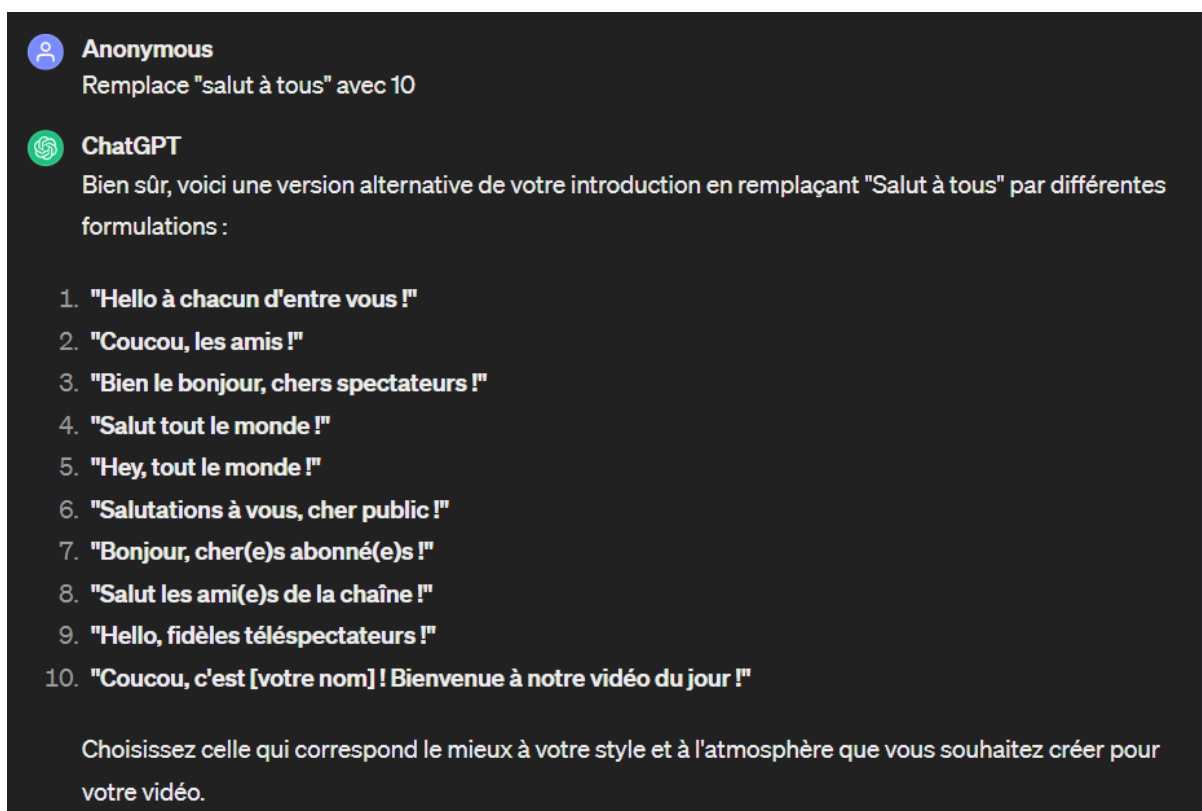


Figure 27 - Une réponse inappropriée où ChatGPT repartage une liste au lieu d'effectuer une modification

La Figure 28 est un autre exemple d'un cas où une requête imprécise a engendré une réponse inappropriée de ChatGPT. Le participant tentait de solliciter une explication des révisions, mais la requête employée manquait de précision, ce qui a incité ChatGPT à repartager le texte révisé au lieu de lister les révisions apportées.

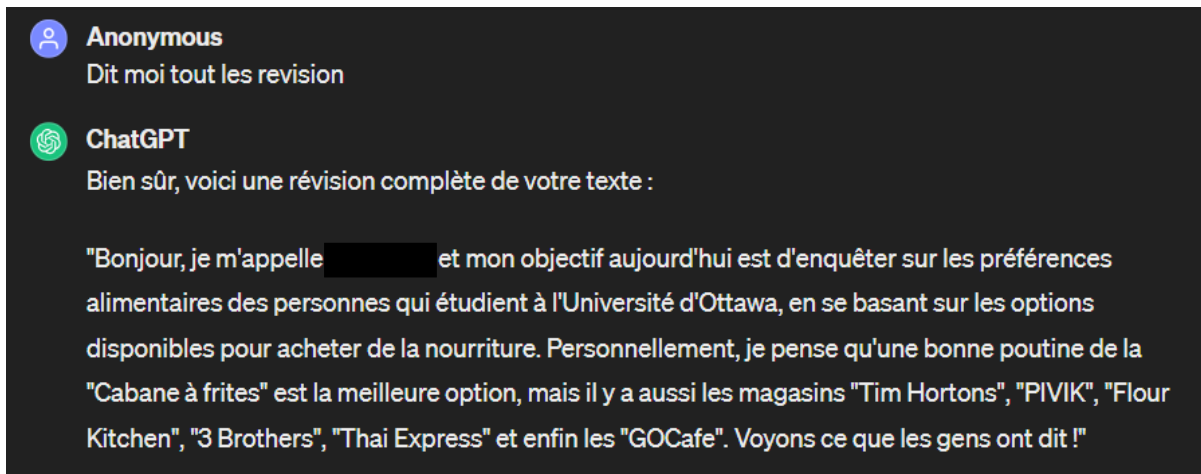


Figure 28 - Une réponse inappropriée où ChatGPT repartage le texte révisé sans inclure d'explications

Au total, 18,2 % des réponses de ChatGPT étaient inappropriées ou contenaient des informations incorrectes. On constate que la qualité des réponses a tendance à souffrir quand une requête est imprécise ou mal formulée.

4.3 Résultats du questionnaire post-intervention

Le questionnaire post-intervention avait pour but de recueillir des données portant sur les perceptions des participants à propos de l'utilisation de ChatGPT pour la révision textuelle. Ce deuxième questionnaire comportait 17 questions, mais tel que mentionné plus haut, on s'est intéressé seulement aux quinze premières questions. Des quinze questions qui figurent dans cette analyse, quatorze étaient des affirmations que les participants ont évaluées en utilisant des échelles de Likert, tout comme pour le premier questionnaire. La quinzième question était une question à réponse courte où les participants ont été invités à lister les requêtes qui ont fourni des rétroactions utiles.

La majorité des participants étaient soit d'accord (50 %) ou plutôt d'accord (31,8 %) d'avoir aimé utiliser ChatGPT pour la révision textuelle. 4,5 % des participants ont été neutres, 9,1 % ont été

plutôt pas d'accord et 4,5 % n'étaient pas d'accord. Le Tableau 8 présente ces résultats pour offrir une vue d'ensemble. La prochaine affirmation portait sur la facilité d'utiliser ChatGPT pour la correction. 68,2 % des participants étaient d'accord qu'utiliser ChatGPT pour la révision textuelle était facile, 22,7 % étaient plutôt d'accord et 9,1 % étaient neutres. Pour l'affirmation « utiliser ChatGPT pour la correction a été intuitif », tous les participants étaient soit d'accord (27,3 %) ou plutôt d'accord (72,7 %).

Réponse	Nombre de réponses	Pourcentage
D'accord	11	50 %
Plutôt d'accord	7	31,8 %
Neutre	1	4,5 %
Pas d'accord	2	9,1 %
Plutôt pas d'accord	1	4,5 %

Tableau 8 - Tableau de synthèse des réponses évaluant l'affirmation « J'ai aimé utiliser ChatGPT pour corriger mon texte »

Un seul participant (4,5 %) a déclaré être plutôt pas d'accord avec l'affirmation « les explications de ChatGPT ont été utiles pour la correction ». Les autres participants étaient tous d'accord (81,8 %) ou plutôt d'accord (13,6 %) avec cet énoncé. Les résultats pour l'affirmation suivante démontrent une tendance similaire. L'énoncé était « les explications de ChatGPT ont été claires » et tous sauf un participant (4,5 %) (qui a déclaré être plutôt pas d'accord) étaient d'accord (40,9 %) ou plutôt d'accord (54,5 %).

La majorité des participants ont aussi déclaré que les explications de ChatGPT étaient toujours (45,5 %) ou souvent (45,5 %) logiques. Un participant (4,5 %) a dit qu'elles n'étaient que parfois logiques et un participant (4,5 %) a indiqué qu'elles n'étaient que rarement logiques. Similairement, 40,9 % des participants ont trouvé les explications de ChatGPT toujours cohérentes et 54,5 % des participants les ont trouvées souvent cohérentes. 4,5 % les ont trouvées rarement cohérentes.

Pour comparer ChatGPT aux autres outils de correction linguistique numérique, on a proposé l'affirmation « utiliser ChatGPT pour la correction a été plus facile qu'utiliser d'autres outils de correction linguistique numériques ». 40,9 % des participants étaient d'accord avec cet énoncé, 27,3 % étaient plutôt d'accord, 22,7 % avaient des sentiments neutres et 9,1 % des participants étaient plutôt pas d'accord.

La prochaine affirmation était « j'avais confiance dans les corrections de ChatGPT ». La majorité des participants étaient soit d'accord (27,3 %) ou plutôt d'accord (59,1 %). 4,5 % étaient neutres et 9,1 % étaient plutôt pas d'accord.

La majorité des participants ont affirmé avoir tiré profit de leurs interactions avec ChatGPT pendant la correction, étant d'accord (40,9 %) ou plutôt d'accord (40,9 %) avec l'énoncé « j'ai bénéficié de mes interactions avec ChatGPT lors de la correction ». 9,1 % des participants étaient neutres et 9,1 % des participants étaient plutôt pas d'accord.

La prochaine question demandait aux participants de répondre seulement s'ils avaient négocié le sens avec ChatGPT. L'énoncé était « en négociant le sens avec ChatGPT, j'ai mieux compris mes erreurs ». Treize participants ont répondu à cette question dont cinq participants (38,5 %) étaient d'accord, cinq (38,5 %) étaient plutôt d'accord, deux (15,4 %) étaient neutres et un participant (7,7 %) était plutôt pas d'accord.

Six participants (27,3 %) étaient d'accord avec l'affirmation « comme outil de correction, ChatGPT compte plus d'affordances que de limites ». Onze participants (50 %) étaient plutôt d'accord avec cet énoncé et cinq participants (22,7 %) ont exprimé des sentiments neutres.

Après l'expérience, 54,5 % des participants ont déclaré des sentiments inchangés envers l'IA. 40,9 % des participants ont indiqué que leur opinion envers l'intelligence artificielle était plus positive qu'auparavant et 4,5 % ont déclaré avoir des sentiments plus négatifs qu'auparavant.

La dernière affirmation à évaluer était « formuler des requêtes pour obtenir de la RCÉ a été facile ». 27,3 % des participants ont déclaré qu'il était toujours facile de formuler des requêtes, 59,1 % ont dit que c'était souvent facile et 9,1 % des participants ont dit que c'était parfois facile. La Figure 29 permet de visualiser ses résultats pour voir en quoi les participants ont trouvé la formulation de requêtes facile.

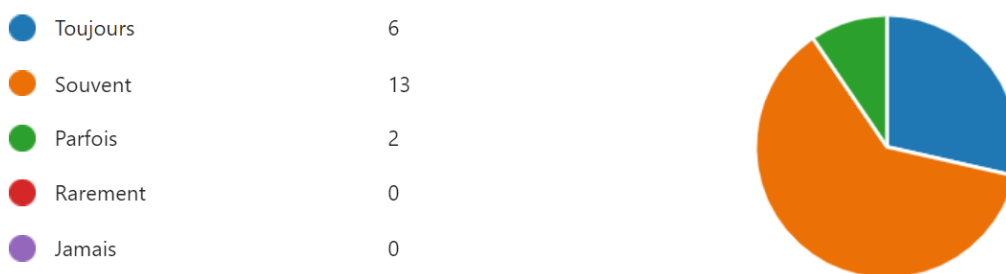


Figure 29 - Diagramme circulaire des réponses à l'affirmation « Formuler des requêtes pour obtenir de la RCÉ a été facile »

Pour la question à réponse ouverte demandant aux participants de lister des exemples de requêtes qui ont fourni des rétroactions utiles, seulement quinze étudiants y ont répondu. Toutefois, certains participants ont listé plusieurs exemples. Les participants ont listé 21 requêtes utiles et trois réponses de ChatGPT utiles.

Des requêtes utiles listées, seize étaient des requêtes dans la catégorie « Clarification ». Des exemples de ce type de requête sont « Je ne comprends pas que.... [sic] », « Pourquoi avez vous? [sic] », « Je ne comprend pas l'explication 2 [sic] », « pourquoi ton structure est mieux? [sic] » et « pourquoi est-ce que "grande" est mieux que "importante"? [sic] ». Cinq des requêtes utiles

identifiées étaient les requêtes imposées ou des variations de celles-ci ; « Réviser ce texte », « Reviser [*sic*] », « corrigez ça », « explique » et « Pouvez-vous expliquer vos changements? ».

Les trois autres réponses à la question étaient les rétroactions de ChatGPT. Ce ne sont pas des requêtes ayant sollicité des rétroactions utiles, mais plutôt des rétroactions de ChatGPT que les participants ont identifiées comme étant utiles. Bien que ces réponses ne correspondent pas aux réponses prévues, il est intéressant de voir les types de rétroactions que les participants ont trouvés utiles. Les trois rétroactions identifiées figuraient toutes dans la catégorie « Réponse attendue » où la machine explique les changements apportés ou des justifications de ceux-ci ; « 1. Maintien de la cohérence 2. Adaptation des références au genre féminin » et « "Est-ce que vous pensez que les frais de universite doivent etre plus petit?" Révision : "Pensez-vous que les frais universitaires devraient être plus bas ?" Explication : J'ai corrigé la formulation pour la rendre plus fluide et grammaticalement correcte. »

Conséquemment, il semble que les participants ont trouvé utiles les rétroactions de ChatGPT quand elles expliquaient les modifications. Ils ont également trouvé utiles les requêtes ayant sollicité ce type de rétroaction.

5 Discussion

Cette section propose une interprétation des résultats présentés dans la section précédente à la lumière des questions de recherche et des écrits recensés. Des recommandations pédagogiques appuyées par les résultats d'autres études sont ensuite suggérées. Les données des deux questionnaires et des fils de discussion analysés ont été croisées afin de faire des inférences. On aborde chaque question de recherche et ensuite on avance des recommandations d'après les résultats obtenus. Enfin, on propose une sous-section portant sur les limites de cette étude ainsi que des suggestions pour de futures études.

5.1 QR1

Quelles sont les pratiques autocorrectives actuelles déclarées par les étudiants dans le cours FLS 2751 concernant la révision textuelle des tâches de rédaction?

Dans le premier questionnaire, la majorité des participants (90,9 %) ont déclaré corriger les textes qu'ils rédigent en français soit souvent, soit toujours. D'ailleurs, ils se servent surtout des outils numériques pour ce faire, les privilégiant à la place des outils traditionnels tels que les dictionnaires et les Bescherelle en papier.

En particulier, ils ont déclaré utiliser Antidote, BonPatron, les logiciels de traitement de texte et d'autres outils non nommés. Il n'est pas surprenant que les participants se penchent sur l'aide des correcticiels Antidote et BonPatron pour corriger leurs textes. Hebbinckuys et collab. (2021) avaient identifié qu'Antidote représente une aide précieuse pour soutenir la rédaction et la révision des textes en français. 18,2 % des participants à notre étude ont déclaré utiliser Antidote et 68,2 % ont déclaré utiliser BonPatron. Étant donné le niveau des étudiants dans le cours (intermédiaire faible), il est logique qu'ils se servent de BonPatron plus qu'Antidote. BonPatron a été conçu pour un public anglophone et emploie moins de métalangage dans ses explications (López Santiago,

2016) tandis qu'Antidote est destiné aux utilisateurs qui maîtrisent déjà le métalangage français (Hebbinckuys, 2021). D'autre part, BonPatron offre une version gratuite, tandis qu'Antidote est payant.

La majorité des participants ont déclaré être plutôt d'accord de comprendre les rétroactions proposées par les outils de correction linguistiques numériques qu'ils exploitent et d'en être satisfaits aussi. En revanche, dans le deuxième questionnaire, la majorité des participants (68,2 %) étaient d'accord ou plutôt d'accord qu'utiliser ChatGPT pour la correction a été plus facile qu'utiliser d'autres outils de correction linguistique numérique, ce qui semble suggérer que les participants n'étaient pas aussi satisfaits des outils qu'ils exploitent qu'ils l'avaient déclaré.

La majorité des participants (86,4 %) ont aussi déclaré être d'accord ou plutôt d'accord d'avoir confiance dans les corrections des outils qu'ils exploitent. D'autre part, dans le deuxième questionnaire, la majorité des participants (86,4 %) ont déclaré être d'accord ou plutôt d'accord d'avoir confiance dans les corrections de ChatGPT. Au total, 19 participants ont déclaré un haut degré de confiance dans les corrections des outils de correction linguistique numériques qu'ils utilisent et 19 participants ont aussi déclaré avoir un haut degré de confiance dans les corrections de ChatGPT. Cela suggère que les participants ont des degrés de confiance comparables pour les corrections des outils explicitement conçus pour la révision textuelle et pour les corrections de ChatGPT. Pendant son étude en 2020, Koltovskaia a trouvé que vers la fin du processus de la révision textuelle, une étudiante d'anglais langue seconde de niveau intermédiaire faible a commencé à accepter les rétroactions proposées par Grammarly sans y réfléchir, ce qui semble démontrer un haut degré de confiance dans les rétroactions proposées. Il n'est donc pas surprenant que la majorité des participants à notre étude (qui sont aussi des apprenants de langue seconde de

niveau intermédiaire faible) aient un haut degré de confiance dans les corrections proposées par les outils numériques.

Parmi les participants à notre étude, l'utilisation des outils d'IA pour la rédaction et pour la correction en français ne semble pas encore très répandue. Cependant, la majorité des participants (54,5 %) ont déclaré des attitudes positives ou plutôt positives envers l'intelligence artificielle, ce qui est conforme aux résultats de l'étude de Shoufan (2023) où la majorité des participants avaient déclaré des sentiments positifs envers ChatGPT.

5.2 QR2

Comment est-ce que les étudiants dans le cours FLS 2751 se sont servis de ChatGPT pour obtenir de la RCÉ afin de réviser leur texte écrit?

On constate que les pratiques de révision textuelle avec ChatGPT sont hétérogènes. Les pratiques autocorrectives des apprenants de L2 sont souvent uniques, tel que démontré par Koltovskaia (2020). En analysant le nombre de tours de paroles, le nombre de requêtes additionnelles et le nombre de mots dans les fils de discussion, il est évident que certains participants se sont engagés plus que d'autres pendant la révision textuelle avec ChatGPT. Les étudiants ayant participé à l'intervention en personne se sont plus investis que ceux ayant participé en ligne probablement parce qu'en ligne, les participants étaient moins encadrés. D'ailleurs, il importe de mentionner que les participants n'ont pas été interrogés sur leurs motivations en tant qu'apprenants du FLS et la motivation a un impact sur l'engagement avec la tâche. Enfin, la tâche n'était pas notée, ce qui aurait pu avoir un impact sur le degré d'engagement des participants (Huang, 2012).

Pendant l'analyse des fils de discussion, on s'est intéressé aux différents types de requêtes formulées par les étudiants. Une typologie comptant sept catégories a été élaborée afin de classer

les requêtes. Ces catégories sont : « Requête imposée », « Clarification », « Socialisation avec la machine », « Désaccord », « Question autre », « Révision » et « Modification ». Les participants ont surtout employé les requêtes imposées (43 occurrences), mais la majorité a aussi inventé de nouvelles requêtes. En particulier, les participants ont formulé des requêtes dans la catégorie « Clarification » (42 occurrences) sollicitant des informations ou des précisions portant sur des éléments dans le texte original, dans le texte révisé ou concernant les deux à la fois. Ces données corroborent ceux de Su et collab. (2023) qui affirment qu'il est possible d'entrer en interaction avec ChatGPT pour solliciter des précisions et des clarifications. Il semble que le fait d'exprimer son incompréhension ou de formuler une demande d'une clarification a permis aux participants de solliciter des rétroactions mieux adaptées à leurs ZPD (Op. cit), ainsi les rendant plus compréhensibles et utiles (Zhang, 2017). C'est le cas, par exemple de P11 qui a reformulé une requête de clarification plusieurs fois afin d'obtenir une réponse qu'il comprenait.

Des 110 requêtes, seulement 20,9 % ne portaient pas sur la tâche. Les requêtes dans la catégorie « Question autre » (3 occurrences) sont un exemple de ce type de requête. La majorité des requêtes inventées par les participants portaient sur la révision textuelle (79,1 %), signifiant qu'ils se sont concentrés sur la tâche linguistique, ce qui semble démontrer un haut degré d'engagement (Egbert, 2020). Par exemple, les requêtes dans la catégorie « Révision » (3 occurrences) sollicitent ou portent sur des révisions additionnelles et les requêtes dans la catégorie « Modification » (2 occurrences) avaient pour but d'éditer le texte. Le fait que certains étudiants ont sollicité de nouvelles révisions et ont essayé de modifier leurs textes de leur propre volonté semble démontrer un haut degré d'engagement.

La nature interactive de ChatGPT (OpenAI, 2024b, Su et collab., 2023) semble avoir accordé aux participants plus d'agentivité et d'autonomie au sein du processus de révision textuelle, ce qui

aurait pu rendre les rétroactions obtenues plus utiles (Zhang, 2017). Une majorité des participants ont posé des questions de suivi portant sur les rétroactions proposées, démontrant de l'agentivité. Selon la définition proposée par Oga-Baldwin, l'agentivité fait référence aux changements apportés par un apprenant pour répondre à ses besoins (2019). Le fait d'avoir posé des questions de suivi et d'avoir remis en question certaines rétroactions proposées semble démontrer que les participants avaient de l'agentivité en interagissant avec ChatGPT pendant la révision textuelle. Cela corrobore aussi l'idée que l'agentivité et l'autonomie promeuvent l'engagement avec la tâche (Mercer, 2019).

Dans le deuxième questionnaire, on a proposé une question à laquelle seuls les étudiants ayant négocié le sens avec ChatGPT devaient répondre. La majorité des treize répondants (76,9 %) ont déclaré être d'accord ou plutôt d'accord que le fait d'avoir négocié le sens avec ChatGPT leur a permis de mieux comprendre leurs erreurs. Les requêtes dans la catégorie « Désaccord » (6 occurrences) sont des exemples de négociation de sens. Selon Johnson (2008), les interactions, en particulier les négociations de sens sont reconnues comme des conditions nécessaires pour l'acquisition d'une langue seconde. Ces négociations de sens aident les apprenants à découvrir ce qu'ils peuvent faire avec une langue, ainsi que ce qu'ils ne peuvent pas faire (Op. cit). Il est donc logique que la majorité des participants aient trouvé qu'en négociant le sens avec ChatGPT, ils ont mieux compris leurs erreurs. D'autre part, selon Nakamura et collab. (2021), les négociations de sens sont des indicateurs verbaux qui peuvent indiquer l'engagement cognitif. Conséquemment, il semble qu'en négociant le sens avec ChatGPT, les participants ont aussi démontré leur engagement cognitif.

On note aussi chez certains participants des degrés élevés d'engagement social (Op. cit), en particulier quand un participant a employé une requête dans la catégorie « Socialisation avec la

machine » (11 occurrences) tels que les salutations, les compliments et les remerciements. Bien que ces requêtes ne soient pas forcément reliées à la révision textuelle, elles semblent humaniser les interactions avec la machine (Cf. Ruan et collab., 2021) les rendant ainsi plus naturelles.

Comme le correcticiel Antidote (Ouellet, 2013), ChatGPT est en mesure de repérer un grand nombre d'erreurs, y compris les erreurs fréquentes des apprenants de FLS (Hamel & Milčević, 2007 ; Mougeon & Mougeon, 2003) et de proposer des corrections pour les rectifier. La RCÉ de ChatGPT est aussi synchrone (immédiate) (Arroyo & Yilmaz, 2018 ; Shintani, 2016 ; Shintani & Aubrey, 2016) et il est possible de resoumettre un texte plusieurs fois pour solliciter de nouvelles révisions et ainsi promouvoir l'autonomie de l'utilisateur (Zhang & Hyland, 2018).

En ce qui concerne les rétroactions proposées par ChatGPT, on constate que la majorité était appropriée et qu'elles partageaient des informations correctes. Une typologie a été élaborée pour classer les réponses de ChatGPT. Cette typologie comprend neuf catégories : « Réponse attendue », « Clarification », « Socialisation avec le participant », « Liste », « Concession », « Modification », « Autocorrection », « Réponse autre » et « Révision ». Les réponses dans la catégorie « Réponse attendue » (43 occurrences) étaient celles répondant aux deux requêtes imposées fournissant une révision initiale ainsi que l'explication de cette dernière. Lors de la révision initiale, ChatGPT a bien révisé les textes soumis, ce qui corrobore les résultats des études de Coyne et collab. (2023), de Fang et collab. (2023) et de Wu et collab. (2023) que cet outil a une bonne capacité pour la correction des erreurs grammaticales, ainsi que les résultats de l'étude de Su et collab. (2023) qu'il est utile pour la révision et la reformulation textuelles. De concert avec Su et collab. (2023), il semble que pour les textes entiers, ChatGPT est en mesure de corriger plusieurs différents types d'erreurs grammaticales et peut modifier les choix lexicaux et les structures syntaxiques afin d'améliorer la concision et le ton professionnel du texte. Il peut également suggérer des

modifications pour améliorer l'organisation textuelle. D'ailleurs, les données de notre étude corroborent ceux de Su et collab. (2023) qui démontrent que ChatGPT peut fournir des commentaires personnalisés, un élément très important pour une RCÉ utile selon Loem et collab. (2023).

On constate que la machine a souvent produit des réponses dans la catégorie « Clarification » (33 occurrences). Ces réponses expliquaient ou clarifiaient des éléments portant sur le texte original, sur le texte révisé ou sur les deux à la fois.

Pendant l'étude de Su et collab. (2023), quand les auteurs ont tenté de faire l'évaluation du contenu d'un texte avec ChatGPT, la machine a fourni des rétroactions évaluatives plutôt que des rétroactions explicatives. Cette tendance n'a pas été répétée pendant notre étude. Ceci est démontré par la Figure 13 précédemment qui illustre une occasion où un participant a sollicité une évaluation concernant son emploi de l'imparfait dans le texte original. ChatGPT a répondu avec une réponse explicative et n'a pas tenté pas d'évaluer le texte.

Su et collab. (2023) ont aussi postulé que ChatGPT peut servir de moteur de recherche permettant de trouver des collocations et d'effectuer des recherches linguistiques. Les résultats de la présente étude corroborent ce postulat. Dans la catégorie « Liste » (7 occurrences), on retrouve des exemples où la machine a fourni des listes d'éléments pour répondre aux recherches linguistiques des utilisateurs. La Figure 12 précédemment est un exemple d'une occasion où un étudiant effectue une recherche linguistique avec ChatGPT pour discerner la différence entre deux mots.

ChatGPT a aussi répondu de manière appropriée hors du cadre de la tâche de révision linguistique. Quand un utilisateur a envoyé une requête de socialisation (11 occurrences), la machine a envoyé une réponse dans la catégorie « Socialisation avec le participant » (11 occurrences). Similairement,

chaque requête dans la catégorie « Question autre » (3 occurrences) a sollicité une réponse dans la catégorie « Réponse autre » (3 occurrences).

En ce qui a trait des défis de ChatGPT pour la RCÉ, on constate qu'il peut fabuler et prétendre avoir effectué des modifications non faites ou mal classer les rétroactions effectuées, bien que le texte soit correctement révisé dans une réponse précédente. Fang et collab. (2023) ont trouvé que ChatGPT a rencontré des difficultés avec certains types d'erreurs, en particulier celles s'étendant au-delà de la délimitation des phrases, durant la CEG en anglais, en allemand et en chinois. Pendant notre étude, ChatGPT n'a pas commis d'erreurs en corrigeant les textes en français pendant la révision initiale, mais on constate qu'il a parfois répété une erreur précédemment corrigée en produisant une nouvelle version du texte. Comme l'ont postulé Fang et collab. (2023), ceci pourrait être en raison des limites de la mémoire contextuelle de ChatGPT.

Wu et collab. (2023) et Fang et collab. (2023) ont trouvé que ChatGPT avait une tendance à surcorriger, ce que certains apprenants pourraient trouver accablant ou décourageant (McMartin-Miller, 2014). Il semble que la majorité des participants à notre étude n'ont pas rencontré ce défi étant donné qu'ils ont déclaré être d'accord ou plutôt d'accord d'avoir aimé utiliser ChatGPT pour corriger leurs textes, ce qui semble démontrer qu'ils n'ont pas été découragés par la quantité de rétroactions.

Des 110 réponses de ChatGPT, vingt (18,2 %) ont été identifiées comme étant incorrectes ou inappropriées. Les requêtes imprécises ou mal formulées ont parfois sollicité de telles réponses, ce qui semble souligner l'importance de bien formuler ses requêtes. Ceci corrobore les résultats obtenus par Coyne et collab. (2023), Fang et collab. (2023) et Loem et collab. (2023) que les requêtes mal formulées ou imprécises peuvent négativement affecter la précision des réponses produites. En écho aux résultats de Loem et collab. (2023), nos données démontrent aussi en quoi

il est important d'inclure suffisamment d'information dans les requêtes afin d'obtenir une réponse utile. La réponse dans la catégorie « Révision » (1 occurrence) et certaines réponses dans la catégorie « Modification » (4 occurrences) n'ont pas été appropriées en raison d'une requête imprécise ou mal formulée. Kojima et collab. (2023) et García-Peñalvo (2023) affirment aussi que le contexte inclus dans la requête affecte la réponse qu'elle sollicite.

En analysant l'alignement requête-réponse, on constate que ChatGPT semble avoir produit des réponses précises et pertinentes aux différentes requêtes, sans employant de réponses scénarisées. Tel que dit plus haut, le fait d'envoyer un type de requête n'a pas toujours obtenu un même type de réponse, ce qui suggère que ChatGPT se sert de l'information dans les requêtes pour répondre de manière précise au lieu de répondre avec des réponses automatiques qui pourraient ne pas être pertinentes en contexte. La machine semble adopter une approche de cas par cas afin de produire des réponses pertinentes aux différentes requêtes et peut conséquemment éviter les réponses scénarisées qui peuvent manquer de pertinence (Belda-Medina & Calvo-Ferrer, 2022). Fang et collab. (2023) ont aussi trouvé que ChatGPT se sert de l'information dans les différentes requêtes pour produire une réponse précise. Dans nos résultats, ceci est illustré par les réponses dans les catégories « Concession » (4 occurrences) et « Autocorrection » (4 occurrences) où la machine ajuste sa réponse en raison de nouvelles informations partagées par l'utilisateur.

5.3 QR3

Après l'intervention, est-ce que les attitudes envers l'IA des étudiants en FLS 2751 ont changé?

La majorité des participants à notre étude n'ont pas changé d'avis concernant l'IA. Des 22 participants, douze (54,5 %) ont déclaré des sentiments inchangés. Cependant, des dix participants ayant changé d'avis, neuf (90 %) ont déclaré des attitudes plus positives envers l'IA et un seul participant (4,5 %) a déclaré avoir une attitude plus négative qu'auparavant. Ces résultats semblent

être en lien avec ceux obtenus par Shoufan (2023) pendant son étude portant sur les perceptions des étudiants envers ChatGPT comme outil pédagogique où les étudiants sondés avaient des perceptions plutôt positives de l'outil.

D'après les résultats du deuxième questionnaire portant sur l'expérience d'avoir utilisé ChatGPT pour la révision textuelle, il semble que ChatGPT est facile et intuitif/convivial à utiliser et qu'il propose des explications majoritairement claires, logiques, cohérentes et utiles pour la correction. D'ailleurs, la majorité des participants étaient d'accord ou plutôt d'accord d'avoir confiance dans les corrections de ChatGPT. Tel que dit plus haut, leur degré de confiance pour les corrections de ChatGPT est comparable à celui pour les corrections des autres outils pour la correction linguistique.

Le deuxième questionnaire ne semble pas avoir relevé de limites flagrantes concernant l'utilisation de ChatGPT pour la RCÉ. La majorité des participants ont exprimé des opinions positives envers l'expérience, bien qu'une petite quantité de participants aient exprimé des sentiments neutres ou plus négatifs concernant certains éléments. Il semble donc que les explications de ChatGPT peuvent occasionnellement manquer de clarté, de logique, de cohérence et d'utilité pour la correction. Cependant, les étudiants ayant exprimé des opinions négatives sont dans la minorité. Ces résultats sont aussi alignés avec ceux obtenus pendant l'étude de Shoufan (2023).

Selon la majorité des participants à notre étude (86,4 %), formuler des requêtes pour obtenir une RCÉ a été toujours ou souvent facile. Seulement deux participants (9,1 %) ont indiqué qu'il n'était que parfois facile et aucun participant n'a déclaré qu'il était rarement ou jamais facile. L'étude de Shoufan (2023) a relevé que la majorité des participants ont trouvé la création de requêtes modérément difficile, mais son étude ne se concentrait pas sur l'utilisation de ChatGPT pour la RCÉ.

La précision des réponses de ChatGPT a obtenu des scores modiques pendant l'étude de Shoufan (2023), mais les participants ont déclaré que poser des questions de suivi a été utile. La majorité des participants à notre étude (86,4 %) ont déclaré être d'accord ou plutôt d'accord d'avoir confiance dans les réponses de ChatGPT avec un participant (4,5 %) étant neutre et deux participants (9,1 %) étant plutôt pas d'accord. Ces résultats suggèrent que les participants étaient plutôt satisfaits de la qualité des réponses de ChatGPT. Des participants ayant négocié le sens avec l'outil, la majorité (76,9 %) était d'accord ou plutôt d'accord que cela les avait aidés à mieux comprendre leurs erreurs, tout comme poser des questions de suivi a été utile pour les participants à l'étude de Shoufan (2023).

On a dit plus haut que les principales limites des robots conversationnels sont qu'ils ont tendance à produire des réponses scénarisées, d'avoir une compréhension limitée, une inefficacité de communication et des réponses illogiques (Belda-Medina & Calvo-Ferrer, 2022). Les résultats de cette étude exploratoire suggèrent que la majorité des participants n'ont pas constaté ces limites chez ChatGPT et que la machine est en mesure de produire des réponses précises et logiques pour répondre aux requêtes précises et clairement formulées. Cependant, les réponses peuvent occasionnellement manquer de pertinence ou de précision.

5.4 Recommandations

Cette sous-section vise à proposer des recommandations pédagogiques à la lumière des résultats obtenus pour optimiser l'utilisation de ChatGPT pour la RCÉ interactive. On s'intéresse d'abord aux recommandations pour la formulation des requêtes et ensuite aux différentes manières d'utiliser ChatGPT pour les tâches de rédaction et de révision textuelle.

On constate que ChatGPT peut souvent répondre avec une réponse correcte et précise, mais peut aussi mécomprendre une requête. En se servant de ChatGPT pour la révision textuelle, afin d'éviter

les incompréhensions de la part de la machine, il est important pour l'apprenant de bien formuler les requêtes et d'être précis. Les requêtes claires et précises permettent un meilleur alignement réponse-requête où la réponse obtenue correspond de manière logique et appropriée à la requête employée. Les requêtes bien formulées respectent les normes grammaticales (Cf. Coyne et collab., 2023) et orthographiques de la langue, emploient un langage simple et sont détaillées et contextualisées. Si la requête porte sur une partie précise du texte, reprendre le verbatim de l'extrait sur lequel on désire solliciter une révision ou une clarification peut permettre d'obtenir une rétroaction plus précise et utile. Contextualiser les requêtes peut aussi aider à éviter des mécompréhensions de la part de la machine. Il est donc recommandé d'indiquer à quelle version du texte à réviser l'on se réfère dans la requête. Conséquemment, une requête comme : « pourquoi est-ce que le mot 'grande' dans le texte original a été remplacé par le mot 'importante' dans le texte révisé? » serait préférable à la requête « pourquoi est-ce que "grande" est mieux que "importante"? [sic] ». Selon Kojima et collab. (2023) ainsi que Fang et collab. (2023), la méthode « Chain of Thought » (CoT), où l'on demande explicitement à la machine de raisonner étape par étape, sollicite des réponses plus claires et précises. Conséquemment, il pourrait être utile à l'enseignant de partager les manières de formuler des requêtes en contexte CoT avec ses apprenants avant de leur faire faire une révision textuelle avec ChatGPT. Si l'on reçoit une réponse insatisfaisante ou incorrecte, il peut être bénéfique d'ajouter de l'information à la requête initiale et ensuite la soumettre de nouveau afin de solliciter une réponse plus appropriée. Les résultats de Loem et collab. (2023) soulignent l'avantage de fournir de l'information additionnelle dans ses requêtes, tel que préciser son niveau de compétence linguistique.

L'analyse des fils de discussion a relevé que parmi les divers types de requêtes formulés par les apprenants, les requêtes qu'ils ont identifiées comme étant les plus utiles étaient celles sollicitant

une révision ou une clarification. On constate aussi que les questions de suivi et les négociations de sens peuvent permettre aux utilisateurs de mieux comprendre leurs fautes et les rétroactions proposées afin d'y remédier. Ainsi, il pourrait être judicieux pour l'enseignant de modéliser une interaction avec ChatGPT pour ses apprenants afin de leur montrer comment bien formuler les divers types de requêtes et comment les modifier pour solliciter des rétroactions mieux adaptées à leurs ZPD. Par exemple, on peut demander à la machine de reformuler une rétroaction précédente en employant un vocabulaire plus simple, en fournissant un exemple concret ou en la traduisant vers une autre langue.

De concert avec Su et collab. (2023), nos données ont démontré qu'il est possible de se servir de ChatGPT pour effectuer des recherches linguistiques, trouver des collocations et trouver des expressions alternatives. L'emploi de ChatGPT comme moteur de recherche pourrait être envisagé au sein d'une tâche de révision textuelle. Su et collab. (2023) ont aussi postulé que ChatGPT pourrait permettre une réflexion sur la tâche linguistique parce que l'outil sauvegarde les fils de discussion. En ayant la possibilité de revoir son fil de discussion après la réalisation d'une tâche, un apprenant pourrait mieux comprendre son processus de rédaction et de révision, ainsi permettant une réflexion post-tâche. Conséquemment, ChatGPT pourrait être intégré en salle de classe non seulement pour la révision textuelle, mais aussi pour développer un regard critique sur le processus rédactionnel après avoir effectué une tâche. Ainsi, nous encourageons les enseignants de FLS à prévoir différentes occasions où ChatGPT pourrait être intégré dans leur enseignement de la rédaction.

Enfin, ChatGPT peut commettre des erreurs, donc il est suggéré de ne pas tenir pour acquis que ses réponses sont toujours correctes. Remettre en question une réponse, négocier le sens ou solliciter une précision additionnelle sont de bonnes stratégies pour vérifier les réponses de la

machine. Il faut exercer son esprit critique pour savoir quand accepter et quand rejeter une rétroaction proposée (Su et collab., 2023). Bien entendu, il faut que les apprenants aient assez de connaissances de la langue pour pouvoir évaluer la qualité des réponses de ChatGPT, ainsi la révision textuelle avec ChatGPT n'est pas une activité qui convient à tous les contextes (Kohnke, Moorhouse & Zou, 2023). La machine peut commettre des fautes, donc il faut apporter un regard critique à ses réponses pour assurer qu'elles sont de qualité. Pour ce faire, on peut utiliser d'autres outils de correction et croiser les rétroactions afin d'en déterminer la qualité. D'ailleurs, une discussion en classe portant sur les avantages et les inconvénients d'utiliser ChatGPT pourrait promouvoir le développement de l'esprit critique de ses apprenants ainsi que leurs compétences en littératie numérique (Godwin-Jones, 2022).

5.5 Les limites de l'étude

Cette étude comporte des limites. La première concerne le nombre restreint de participants : 22 au total, issus d'un même cours de FLS, ce qui constitue un petit échantillon. Afin d'obtenir des données qui peuvent être généralisées de manière plus fiable, de futures études pourraient être menées avec un plus grand échantillon de participants, de différents niveaux de compétence linguistique. D'autre part, cette étude portait sur une seule tâche de production écrite. De futures études pourraient être menées avec plusieurs types de tâches d'apprentissage pour explorer l'utilité de la RCÉ de ChatGPT dans divers contextes.

En raison de contraintes temporelles, on a dû distribuer le deuxième questionnaire immédiatement après l'intervention ponctuelle. Ceci n'était pas originalement planifié et a limité le temps de réflexion des participants après la tâche. D'ailleurs, en raison de la distribution du questionnaire si rapide après l'intervention, on n'a pas pu obtenir des données portant sur si les apprenants avaient continué à utiliser ChatGPT pour la révision textuelle. Enfin, la chercheuse a été présente pendant

l'intervention et pendant que les participants ont répondu aux questionnaires. Sa présence pendant l'intervention aurait pu inciter les participants à être plus ou moins engagés avec la tâche et le fait que les participants savaient qu'elle (la personne ayant mené l'intervention) allait être la personne à lire leurs réponses aux questionnaires pourrait aussi avoir influé sur leurs réponses.

Cette étude, étant de nature exploratoire, a nécessité un certain degré de subjectivité lors de l'étape d'analyse. En adoptant une approche post-positiviste (Panhwar, Ansari & Shah, 2017), nous avons essayé d'équilibrer l'analyse subjective et à la fois objective des données récoltées. D'ailleurs, l'étude avait surtout pour but d'analyser l'utilité des rétroactions de ChatGPT pour les étudiants de FLS en contexte de révision textuelle. Il serait bénéfique pour de futures études d'évaluer non seulement l'utilité des réponses de ChatGPT pour la révision textuelle en français, mais également la qualité de ces dernières par le biais d'une métrique quantitative. Un score F_1 ou F_{05} (Fang et collab., 2023) pourrait être employé en tandem avec une analyse qualitative. Dans le domaine du traitement du langage naturel, ces scores évaluent la précision et le rappel d'un système. Ainsi, une telle métrique est utile pour mesurer en quoi un système (tel que ChatGPT) peut effectuer une variété de tâches linguistiques, y compris la correction des erreurs grammaticales.

Enfin, pour une future étude, il serait intéressant de demander aux participants de produire une version définitive de leur texte afin de voir les rétroactions de ChatGPT qui ont été conservées, ainsi que celles qui ont été rejetées. De cette manière, on pourrait voir plus clairement si les rétroactions et les interactions avec ChatGPT permettent aux apprenants de mieux comprendre leurs fautes afin de produire un texte de meilleure qualité. On pourrait également les interroger sur le processus d'interaction avec ChatGPT pour solliciter une rétroaction.

6 Conclusion

Cette étude empirique exploratoire a permis d'offrir un premier aperçu de l'utilisation de ChatGPT pour la révision textuelle interactive en apprentissage du français langue seconde. La RCÉ occupe une place importante dans l'apprentissage des langues secondes (Woodworth & Barkaoui, 2020), mais il peut être lourd (voire difficile) pour les enseignants de fournir des rétroactions utiles à leurs apprenants dans des délais raisonnables (Op. cit) et la majorité des outils numériques pour la RCÉ permettent peu (ou pas) d'interaction au sein du processus de la révision textuelle (Op. cit).

On s'est intéressé à ChatGPT pour la RCÉ en raison de sa nature interactive (OpenAI, 2024b ; Su et collab., 2023) et sa capacité pour la CEG (Coyné et collab., 2023 ; Fang et collab., 2023 ; Wu et collab., 2023). À notre connaissance, cette étude est la première à explorer l'utilisation de ChatGPT pour la RCÉ en français. Il faudrait davantage de recherches pour déterminer de manière fiable que ChatGPT est un outil utile pour la révision textuelle interactive. Cela dit, les résultats obtenus suggèrent que ChatGPT pourrait être un outil utile pour fournir une RCÉ interactive aux apprenants de FLS de niveau intermédiaire faible. ChatGPT semble adopter une démarche de cas par cas afin de répondre aux requêtes des utilisateurs et ainsi la machine évite de répondre en employant des réponses scénarisées ou automatiques qui manquent souvent de pertinence (Belda-Medina & Calvo-Ferrer, 2022). Toutefois, pour obtenir les réponses les plus pertinentes et utiles, il faut employer des requêtes précises et bien formulées (Fang et collab., 2023 ; Loem et collab., 2023) et apporter un regard critique aux réponses obtenues (Su et collab., 2023).

Cette étude a permis de récolter un corpus de fils de discussion d'interactions apprenants-machine et d'élaborer une typologie pour classer les requêtes formulées par les étudiants. Cette typologie donne une vue d'ensemble des différentes requêtes et permet une meilleure compréhension des types de rétroactions que les apprenants souhaitent obtenir. Une deuxième typologie a été conçue

pour classer les réponses de ChatGPT, ce qui a permis de voir les manières dont la machine répond aux différentes requêtes. En analysant quels types de requêtes ont sollicité quels types de réponses, on a pu formuler des recommandations portant sur la création des requêtes afin d'aider les utilisateurs à obtenir des réponses pertinentes et précises.

Enfin, les participants à cette étude ont exprimé des opinions majoritairement favorables concernant l'utilisation de ChatGPT pour la révision textuelle et semblent avoir tiré profit de leurs interactions avec la machine pour obtenir une RCÉ. D'après les résultats, les principales affordances de ChatGPT pour la révision textuelle sont le fait qu'il est possible de négocier le sens et d'obtenir des explications additionnelles pendant l'interaction. Ces affordances semblent promouvoir l'agentivité et l'autonomie des apprenants et conséquemment leur engagement aussi, ce qui pourrait rendre la RCÉ sollicitée plus utile (Op. cit).

La présente étude compte parmi les recherches pionnières portant sur l'utilisation de ChatGPT pour la révision textuelle. Dans le cadre de cette étude, on s'est concentré sur la RCÉ en contexte d'apprentissage du français langue seconde, un contexte qui, à notre connaissance, demeure relativement sous-exploré. Il faudrait plus de recherches dans ce domaine, mais les données de cette étude sont prometteuses et semblent démontrer que ChatGPT a des forces pour la révision textuelle en français, surtout en raison de sa capacité pour la correction des erreurs grammaticales et sa nature interactive.

7 Références

- Arroyo, D. C. & Yilmaz, Y. (2018). An Open for Replication Study: The Role of Feedback Timing in Synchronous Computer-Mediated Communication [Une étude ouverte à la réplication : Le rôle du temps pour la rétroaction dans la communication synchrone assistée par ordinateur]. *Language Learning*, 68(4), 942-972. doi : 10.1111/lang.12300
- Bandura, A. (2001). Social Cognitive Theory: An Agentic Perspective [La théorie sociocognitive : une perspective agentique]. *Annual Review of Psychology*, 52(1), 1-26. <https://doi.org/10.1146/annurev.psych.52.1.1>
- Bang, Y., Cahyawijaya, S., Lee, N., Dai, W., Su, D. & Wilie, B., ... Fung, P. (2023). *A Multitask, Multilingual, Multimodal Evaluation of ChatGPT on Reasoning, Hallucination, and Interactivity* [Une évaluation multitâche, multilingue et multimodale de ChatGPT sur le raisonnement, l'hallucination et l'interactivité]. <https://doi.org/10.48550/arXiv.2302.04023>
- Baralt, M., Gurzynski-Weiss, L. & Kim, Y. (2016). Engagement with the language: How examining learners' affective and social engagement explains successful learner generated attention to form. Dans M. Sato, & S. Ballinger (dir.), *Peer Intersection and Second Language Learning: Pedagogical potential and research agenda* (p. 209-240). doi : 10.1075/llt.45
- Barbin, F., (2020). La traduction automatique neuronale, un nouveau tournant ?. *Palimpeste*, 4, 51-53. Tiré de <https://shs.hal.science/halshs-03603588>
- Barkaoui, K. (2007). Teaching Writing to Second Language Learners: Insights from Theory and Research. *TESL Reporter* 40(1), 35-48. Tiré de <https://www.researchgate.net/file.PostFileLoader.html?id=55bd22735f7f71e9268b456e&assetKey=AS%3A273823486676992%401442296014069>
- Belda-Medina, J. & Calvo-Ferrer, J. R. (2022). Using Chatbots as AI Conversational Partners in Language Learning [Utiliser les robots conversationnels comme partenaires conversationnels de l'IA en apprentissage des langues]. *Applied Science*, 12(17). <https://doi.org/10.3390/app12178427>
- Bibauw, S., François, T. & Desmet, P. (2022). Dialogue Systems for Language Learning: Chatbots and Beyond. Dans N. Ziegler & M. González-Lloret (dir.), *The Routledge Handbook of Second Language Acquisition and Technology* (1er éd., p. 121-135). <https://doi.org/10.4324/9781351117586>
- Bilmes, J. (2009). Taxonomies are for talking: Reanalyzing a Sacks classic. *Journal of Pragmatics*, 41(8), 1600-1610. doi : 10.1016/j.pragma.2008.10.008
- Bitchener, J. & Ferris, D. R. (2011). *Written Corrective Feedback in Second Language Acquisition and Writing* [La correction rétroactive écrite dans l'acquisition et l'écriture

en langue seconde] [Livre numérique]. <https://doi-org.proxy.bib.uottawa.ca/10.4324/9780203832400>

- Briggs, N. (2018). Neural machine translation tools in the language learning classroom: Students' use, perceptions, and analyses. *The JALT CALL Journal*, 14(1), 3-24. Tiré de <https://web.p.ebscohost.com/ehost/pdfviewer/pdfviewer?vid=0&sid=f9ad2335-f616-40a1-bfef-4c019b280fc2%40redis>
- Buckingham, L., & Aktuğ-Ekinci, D. (2017). Interpreting coded feedback on writing: Turkish EFL students' approaches to revision. *Journal of English for Academic Purposes*, 26, 1-16. <https://doi.org/10.1016/j.jeap.2017.01.001>
- Chang, C., Cunningham, K. J., Müge Satar, H. & Strobl, C. (2018). Electronic feedback on second language writing: A retrospective and prospective essay on multimodality. *Writing and Pedagogy*, 9(3), 405-428. <https://doi.org/10.1558/wap.32515>
- Christenson, S. L. & Reschly, A. L. (2022). Jingle, jangle, and conceptual haziness: Evolution and future directions of the engagement construct. Dans S. Christenson, A. Reschly, & C. Wylie (dir.), *Handbook of Research on Student Engagement* (2e éd., p. 3-19). Springer International Publishing. https://doi.org/10.1007/978-1-4614-2018-7_1
- Coyne, S., Sakaguchi, K., Galvan-Sosa, D., Zock, M. & Inui, K. (2023). *Analyzing the Performance of GPT-3.5 and GPT-4 in Grammatical Error Correction*. Tiré de <https://arxiv.org/abs/2303.14342>
- Creely, E. (2023). The possibilities, limitations, and dangers of generative AI in language learning and literacy practices. *Proceedings of the International Graduate Research Symposium 2023*, 1–14. University of Languages and International Studies (ULIS). Tiré de <https://www.researchgate.net/publication/375558249>
- Druide informatique inc. (1996). Antidote (version 9) [logiciel]. Druide Informatique. <https://www.antidote.info/fr/>
- Druide informatique inc. (2024). A Reformulated Antidote—Now With Generative AI! Tiré de <https://www.antidote.info/en/blog/news/a-reformulated-antidote-now-with-generative-ai>
- Egbert, J. (2020). Language Task Engagement (Theory). Dans J. Egbert & M. F. Roe (dir.), *Theoretical Models for Teaching and Research*. Tiré de <https://opentext.wsu.edu/theoreticalmodelsforteachingandresearch/chapter/language-task-engagement-theory/>
- El Ebyary, K. & Windeatt, S. (2010). The Impact of Computer-Based Feedback on Students' Written Work [L'impact de la rétroaction numérique sur la production écrite des étudiants]. *International Journal of English Studies*, 10(2). <https://doi.org/10.6018/ijes/2010/2/119231>

- Ellis, F. (2009). A Typology of Written Corrective Feedback Types. *ELT Journal*, 63(2), 97-107.
doi : <https://doi.org/10.1093/elt/ccn023>
- Ellis, R., Sheen, Y., Murakami, M., & Takashima, H. (2008). The Effects of Focused and Unfocused Written Corrective Feedback in an English as a Foreign Language Context [Les effets de la rétroaction corrective écrite concentrée et non concentrée dans un contexte d'apprentissage d'anglais langue étrangère]. *System*, 36, 353-371.
<https://doi.org/10.1016/j.system.2008.02.001>
- Fang, T., Yang, S., Lan, K., Wong, D. F., Hu, J., Chao, L. S. & Zhang, Y. (2023). Is ChatGPT a Highly Fluent Grammatical Error Correction System? A Comprehensive Evaluation [Est-ce que ChatGPT est un système de correction d'erreurs grammaticales très courant? Une évaluation compréhensive]. ArXiv. <https://doi.org/10.48550/arXiv.2304.01746>
- Ferris, D. R. (2003). *Response to Student Writing : Implications for Second-Language Students*. Lawrence Erlbaum Associates. Tiré de <https://smartlib.umri.ac.id/assets/uploads/files/99a38-responds-to-students-writing.pdf>
- Ferris, D. R. (2010). Second language writing research and written corrective feedback in SLA: Intersections and practical applications. *Studies in Second Language Acquisition*, 32(2), 181-201. <https://doi.org/10.1017/S0272263109990490>
- Fredholm, K. (2014). Effects of online translation on morphosyntactic and lexical-pragmatic accuracy in essay writing in Spanish as a foreign language. Dans S. Jager, L. Bradley, E. J. Meima, & S. Thouësny (dir.), *CALL Design: Principles and Practice; Proceedings of the 2014 EUROCALL Conference, Groningen, Pays-Bas* (p. 96-101). Research-publishing.net. doi : 10.14705/rpnet.2014.000201
- Fredholm, K. (2015). Online translation use in Spanish as a foreign language essay writing: Effects on fluency, complexity and accuracy. *Revista Nebrija de Lingüística Aplicada a la Enseñanza de las Lenguas*, 18, 1–18. Tiré de <https://www.nebrija.com/revista-linguistica/online-translation-use-in-spanish-as-a-foreign-language-essay-writing-effects-on-fluency-complexity-and-accuracy.html>
- Fredholm, K. (2019). Effects of Google translate on lexical diversity: Vocabulary development among learners of Spanish as a foreign language. *Revista Nebrija de Lingüística Aplicada a la Enseñanza de las Lenguas*, 13(26), 98–117. Tiré de <https://revistas.nebrija.com/revista-linguistica/article/view/300>
- Galaczi, E. (2023). *English language education in the era of generative AI: our perspective*. Cambridge University Press and Assessment. Tiré de <https://www.cambridgeenglish.org/Images/685411/english-language-education-in-the-era-of-generative-ai-our-perspective.pdf>

- García-Peñalvo, F. J. (2023). The perception of Artificial Intelligence in educational contexts after the launch of ChatGPT: Disruption or Panic?. *Education in the Knowledge Society*, 24, 1-9. <https://doi.org/10.14201/eks.31279>
- Godwin-Jones, R. (2022). Partnering with AI: Intelligent writing assistance and instructed language learning [Un partenariat avec l'IA : L'assistance à la rédaction intelligente et l'apprentissage des langues instruit]. *Language Learning and Technology*, 26(2), 5-24. <http://doi.org/10125/73474>
- Google. (2024). Google traduction. Tiré de <https://translate.google.com/>
- Grammarly. (2024). *Introducing generative AI assistance*. Tiré de <https://support.grammarly.com/hc/en-us/articles/14528857014285-Introducing-generative-AI-assistance>
- Guénette, D. (2009). The Cyberscript Project: A mixed-method study of pre-service ESL teachers' corrective feedback beliefs and practices [Thèse de doctorat, Université McGill]. ProQuest Dissertations Publishing. Tiré de <https://www.proquest.com/docview/2812056718?%20Theses&parentSessionId=5zpFP4UBiNHOpN8zOQk3T2%2BPDgQvRx%2FC9PLCifyHNIM%3D&parentSessionId=WkjD5bEKkku%2Bieq70liyHDVzL8wpDjw2UvJbTTSpRe8%3D&pq-origsite=primo&sourcetype=Dissertations%20>
- Guo, K., Wang, J. & Chu, S. K. W. (2022). Using chatbots to scaffold EFL students' argumentative writing. *Assessing Writing* 54, 1-6. <https://doi.org/10.1016/j.asw.2022.100666>
- Hamel, M-J. & Bibeau, L-D. (2021). Pratiques (numériques) de rétroaction corrective des enseignants de langue et prototype d'outil numérique pour les optimiser. *Apprentissage des langues et systèmes d'information et de communication*, 24(2). <https://doi.org/10.4000/alsic.5550>
- Hamel, M-J. & Milčević, J. (2007). Analyse d'erreurs lexicales d'apprenants du FLS : démarche empirique pour l'élaboration d'un dictionnaire d'apprentissage. *Canadian Journal of Applied Linguistics*, 10(1), 25-45. Tiré de <https://journals.lib.unb.ca/index.php/CJAL/article/view/19733>
- Hebbinckuys, N., Hong, R. & Lory, M-P. (2021). Une plateforme interactive autonomisante pour favoriser la réflexion des apprenant.e.s en production écrite : présentation d'un projet-pilote collaboratif et interdisciplinaire autour d'Antidote. *Nouvelle Revue Synergies Canada*, 14, 1-11. Tiré de <https://id.erudit.org/iderudit/1084952ar>
- Heift, T. (2010). Developing an Intelligent Language Tutor [Développer un tuteur de langue intelligent]. *CALICO Journal* 27(3), 443-459. Tiré de <https://www.jstor.org/stable/10.2307/calicojournal.27.3.443>

- Heift, T. (2016). Learner personas and the effects of instructional scaffolding on working behaviour and linguistic performance. Dans C. Caws & M-J. Hamel (dir.), *Language-Learner Computer Interactions: Theory, methodology and CALL applications* (vol. 2, p. 117-136). doi : 10.1075/lse.2
- Huang, S-C. (2012). Pushing Learners to Work Through Tests and Marks: Motivating or Demotivating? A Case in a Taiwanese University. *Language Assessment Quarterly*, 9(1), 60-77. <https://doi.org/10.1080/15434303.2010.510898>
- IBM. (2024). Qu'est-ce que l'apprentissage en profondeur?. Tiré de <https://www.ibm.com/fr-fr/topics/deep-learning>
- Ji, H., Han, I. & Ko, Y. (2023). A systematic review of conversational AI in language education: focusing on the collaboration with human teachers [Une revue systématique de l'IA conversationnel en éducation des langues : concentrant sur la collaboration avec les enseignants humains]. *Journal of Research on Technology in Education*, 55(1), 48-63. doi : 10.1080/15391523.2022.2142873
- Jia, F., Sun, D., Ma, Q & Looi, C-K. (2022). Developing an AI-Based Learning System for L2 Learners' Authentic and Ubiquitous Learning in English Language. *Sustainability* 2022, 14(23). <https://doi.org/10.3390/su142315527>
- Jiao, W., Wang, W., Huang, J-T., Wang, X., Shi, S. & Tu, Z. (2023). *Is ChatGPT a Good Translator? Yes With GPT-4 as the Engine*. ArXiv. <https://doi.org/10.48550/arXiv.2301.08745>
- Johnson, M. (2008). A Philosophy of Second Language Acquisition [Livre numérique]. Tiré de <https://www-degruyter-com.proxy.bib.uottawa.ca/document/doi/10.12987/9780300129410/pdf>
- Jolley, J. R. & Maimone, L. (2015). Free Online Machine Translation: Use and Perceptions by Spanish Students and Instructors. Dans A. J. Moeller (dir.), *Learn languages, explore cultures, transform lives : 2015 Report of the Central States Conference on the Teaching of Foreign Languages* (p. 181-200). Robert M. Terry. Tiré de https://www.academia.edu/11355802/Free_Online_Machine_Translation_Use_and_Perceptions_by_Spanish_Students_and_Instructors
- Kaplan-Rakowski, R., Grotewold, K., Hartwick, P. & Papin, K. (2023). Generative AI and Teachers' Perspectives on Its Implementation in Education [L'IA générative et les perspectives des enseignants sur son intégration en éducation]. *Journal of Interactive Learning Research* 34(2), 313-338. Tiré de https://www.academia.edu/106455593/Generative_AI_and_Teachers_Perspectives_on_Its_Implementation_in_Education
- Kohnke, L., Moorhouse, B. L., Zou, D. (2023). ChatGPT for Language Learning and Teaching. *RELC Journal* 54(2), 537-550. <https://doi.org/10.1177/00336882231162868>

- Kojima, T., Gu, S. S., Reid, M., Matsuo, Y. & Iwasawa, Y. (2023). *Large Language Models are Zero-Shot Reasoners*. Proceedings of the 36th Conference on Neural Information Processing Systems, *NeurIPS 2023*, 1-15. <https://doi.org/10.48550/arXiv.2205.11916>
- Koltovskaia, S. (2020). Student engagement with automated written corrective feedback (AWCF) provided by Grammarly: A multiple case study. *Assessing Writing* 44. doi : 10.1016/j.asw.2020.100450
- Kushmar, L.V., Vornachev, A.O Korobova. I.O., & Kaida, N.O. (2022). Artificial Intelligence in Language Learning: What Are We Afraid of. *Arab World English Journal (AWEJ) Special Issue on CALL*, 8, 262-273. <https://dx.doi.org/10.24093/awej/call8.18>
- Kwon, S. K., Shin, D. & Lee, Y. (2023). The application of chatbot as an L2 writing practice tool. *Language Learning and Technology* 27(1), 1-19. <https://doi.org/10.125/73541>
- Lee, I. (2003). L2 writing teachers' perspectives, practices and problems regarding error feedback. *Assessing Writing*, 8, 216-237. doi :10.1016/j.asw.2003.08.002
- Lee, I (2017). Classroom Writing Assessment and Feedback in L2 School Contexts [Livre numérique]. Tiré de <https://books-scholarsportal-info.proxy.bib.uottawa.ca/en/read?id=/ebooks/ebooks3/springer/2017-08-17/1/9789811039249#page=6>
- Lee, I. (2020). Utility of focused/comprehensive written corrective feedback research for authentic L2 writing classrooms. *Journal of Second Language Writing*, 49. doi : 10.1016/j.jslw.2020.100734
- Levy, M. & Caws, C. (2016). CALL design and research Taking a micro and macro view. Dans C. Caws & M-J. Hamel (dir.), *Language-Learner Computer Interactions: Theory, methodology and CALL applications* (vol. 2, p. 89-113). doi : 10.1075/lse.2
- Lim, F. V. & Phua, J. (2019). Teaching writing with language feedback technology. *Computers and Composition* 54. <https://doi.org/10.1016/j.compcom.2019.102518>
- Loem, M., Kaneko, M., Takase, S. & Okazaki, N. (2023). *Exploring Effectiveness of GPT-3 in Grammatical Error Correction: A Study on Performance and Controllability in Prompt-Based Methods*. Proceedings of the 18th Workshop on Innovative Use of NLP for Building Educational Applications (BEA 2023), 205-219. <https://doi.org/10.48550/arXiv.2305.18156>
- López Santiago, M. (2016). BonPatron.com. *The EUROCALL Review*, 24(2), 51-55. doi : <https://doi.org/10.4995/eurocall.2016.6515>
- Louis-Charles, J-A. (2012). *Les erreurs récurrentes en français, langue seconde (FLS) : un code de correction comme outil de référence peut-il aider les apprenants à les corriger?* [Mémoire de maîtrise, Université du Québec à Montréal]. Archipel. Tiré de <https://archipel.uqam.ca/5094/1/M12636.pdf>

- Luyat, M. & Regia-Corte, T. Les affordances : de James Jerome Gibson aux formalisations récentes du concept. *L'Année psychologique*, 109, 297-332. <https://doi.org/10.3917/anpsy.092.0297>
- McMartin-Miller, C. (2014). How much feedback is enough?: Instructor practices and student attitudes toward error treatment in second language writing. *Assessing Writing*, 19, 24-35. <https://doi.org/10.1016/j.asw.2013.11.003>
- Mercer, S. (2019). *Language Learner Engagement: Setting the Scene [Engagement des apprenants de langue : mise en scène]*. Dans Gao, X (dir.), *Second Handbook of English Language Teaching* (p. 643-660). Tiré de https://link.springer.com/content/pdf/10.1007/978-3-030-02899-2_40.pdf
- Mireault, M-H. (2009). *L'apport des correcticiels pour la correction de textes d'élèves du secondaire* [Mémoire de maîtrise, Université de Montréal]. Papyrus. <https://doi.org/1866/8109>
- Mougeon, R. & Mougeon, F. (2003). *Acquisition et enseignement du français dans les programmes d'immersion*. Tiré de <http://www.yorku.ca/fmougeon/documents/GuideMilleImages.doc>
- Moulieswaran, N. & Prasantha Kumar, N. S. (2023). Investigating ESL Learners' Perception and Problem towards Artificial Intelligence (AI) -Assisted English Language Learning and Teaching [Enquête sur la perception et les problèmes des apprenants d'anglais langue seconde à l'égard de l'apprentissage et de l'enseignement de l'anglais assistés par l'intelligence artificielle (IA)]. *World Journal of English Language*, 13(5), 290-298. doi : 10.5430/wjel.v13n5p290
- Murray, T. (1999). Authoring Intelligent Tutoring Systems: An Analysis of the State of the Art [Création de systèmes de tutorat intelligents : Une analyse de l'état de l'art]. *International Journal of Artificial Intelligence in Education*, 10, 98-129. Tiré de https://www.researchgate.net/publication/32229467_Authoring_Intelligent_Tutoring_Systems_An_Analysis_of_the_State_of_the_Art
- Nakamura, S., Phung, L. & Reinders, H. (2021). The Effect of Learner Choice on L2 Task Engagement. *Studies in Second Language Acquisition*, 43, 428-441. <https://doi.org/10.1017/S027226312000042X>
- Nazari, N., Shabbir, M. S. & Setiawan, R. (2021). Application of Artificial Intelligence powered digital writing assistant in higher education: randomized controlled trial. *Heliyon*, 7(5). doi : 10.1016/j.heliyon.2021.e07014
- Office québécois de la langue française. (2023). Grand dictionnaire terminologique. Tiré de <https://vitrinelinguistique.oqlf.gouv.qc.ca/fiche-gdt/fiche/26570880/grand-modele-de-langage>

- Oga-Baldwin, W. L. Q. Acting, thinking, feeling, making, collaborating: The engagement process in foreign language learning [Agir, penser, sentir, faire, collaborer : le processus d'engagement en apprentissage des langues étrangères]. *System*, 86. <https://doi.org/10.1016/j.system.2019.102128>
- O'Neill, E. M. (2016). Measuring the impact of online translation on FL writing scores. *The International Association for Language Learning Technology*, 46(2), 1-39. <https://doi.org/10.17161/iallt.v46i2>
- O'Neill, E. M. (2019). Training students to use online translators and dictionaries: The impact on second language writing scores. *International Journal of Research Studies in Language Learning*, 8(2), 47-65. <https://doi.org/10.5861/ijrsl.2019.4002>
- OpenAI. (2024a). ChatGPT. Tiré de <https://chat.openai.com/>
- OpenAI. (2024b). Introducing ChatGPT [Introduction de ChatGPT]. Tiré de <https://openai.com/blog/chatgpt>
- OpenAI. (2024c). Pioneering Research on the Path to AGI [Être pionnier dans la recherche pour l'IAG]. Tiré de <https://openai.com/research/overview>
- Ouellet, M. (2013). *Mesure et évaluation des apports d'un correcticiel*. Tiré de <http://www.cdc.qc.ca/parea/788538-ouellet-mesure-evaluation-apports-correcticiel-drummondville-PAREA-2013.pdf>
- Panhwar, A. H., Ansari, S. & Shah, A. A. (2017). Post-positivism: an effective paradigm for social and educational research. *International Research Journal of Arts & Humanities*, 45(45), 253–259. Tiré de https://www.proquest.com/docview/2044301228?rfr_id=info%3Axri%2Fsid%3Aprimo&parentSessionId=F%2B701Vi%2BEUiM4xu%2BEYFCWaISvPq%2BnywPDanW4Uu49xg%3D
- Poibeau, T. (2017). *Machine Translation* [livre numérique]. Tiré de <https://direct-mit-edu.proxy.bib.uottawa.ca/books/book/3127/Machine-Translation>
- Ranalli, J. (2018). Automated written corrective feedback: how well can students make use of it?. *Computer Assisted Language Learning* 31(7), 653-674. <https://doi.org/10.1080/09588221.2018.1428994>
- Ranalli, J., Link, S. & Chukharev-Hudilainen, E. (2017). Automated writing evaluation for formative assessment of second language writing: investigating the accuracy and usefulness of feedback as part of argument-based validation. *Educational Psychology (Dorchester-on-Thames)*, 37(1), 8-25. <https://doi.org/10.1080/01443410.2015.1136407>

- Reeve, J. (2012). A self-determination theory perspective on student engagement. Dans Reschly, A. L., Christenson, S. L. & C. Wylie (dir.) *Handbook of Research on Student Engagement* (2e éd., p. 149-172). doi : 10.1007/978-3-031-07853-8
- Reinders, H. & Mohebbi, H. (2018). Written Corrective Feedback: The Road Ahead [La rétroaction corrective écrite : le chemin à parcourir]. *Language Teaching Research Quarterly*, 6, 1-6. <https://doi.org/10.32038/ltrq.2018.06.01>
- Rhéaume, M. (2023). Plan de cours. [Site Web du cours FLS 2751]. Tiré de <https://uottawa.brightspace.com/d21/le/content/319995/Home>
- Rossoni, L. & ChatGPT. (2022). A inteligência artificial e eu: escrevendo o editorial juntamente com o ChatGPT. *Revista Eletrônica de Ciência Administrativa*, 21(3), 399-405. doi : <http://dx.doi.org/10.21529/RECADM.2022ed3>
- Ruan, S., Jiang, L., Xu, Q., Davis, G. M., Liu, Z., Brunskill, E. & Landay, J. A. (2021). EnglishBot: An AI-Powered Conversational System for Second Language Learning. *26th International Conference on Intelligent User Interfaces*, 26, 434-444. <https://doi.org/10.1145/3397481.3450648>
- Şakrak-Ekin, G. & Balçikanlı, C. (2019). Written Corrective Feedback: EFL Teachers' Beliefs and Practices. *The Reading Matrix: An International Online Journal*, 19(1), 114-128. Tiré de <https://www.readingmatrix.com/files/20-47d49p9h.pdf>
- Scarcella, R. C. (1996). Secondary Education in California and Second Language Research: Instructing ESL Students in the 1990s. *CATESOL Journal* 9(1), 129-152. Tiré de http://www.catesoljournal.org/wp-content/uploads/2016/11/CJ9_scarcella.pdf
- Selvi, A. F. (2019). Qualitative Content Analysis. Dans *The Routledge Handbook of Research Methods in Applied Linguistics* (440-452). Routledge. <https://www-taylorfrancis-com.proxy.bib.uottawa.ca/chapters/edit/10.4324/9780367824471-37/qualitative-content-analysis-ali-fuad-selvi?context=ubx&refId=76aac14b-e5e7-447e-9796-cf9032e95e3e>
- Shintani, N. (2016). The effects of computer-mediated synchronous and asynchronous direct corrective feedback on writing: a case study. *Computer Assisted Language Learning*, 29(3), 517-538. doi : 10.1080/09588221.2014.993400
- Shintani, N. & Aubrey, S. (2016). The Effectiveness of Synchronous and Asynchronous Written Corrective Feedback on Grammatical Accuracy in a Computer-Mediated Environment [L'efficacité de la rétroaction corrective écrite synchrone et asynchrone sur la précision grammaticale dans un environnement informatisé]. *The Modern Language Journal*, 100(1), 296-319. doi : 10.1111/modl.12317

- Shintani, N., Ellis, R. & Suzuki, W. (2014). Effects of written feedback and revision on learners' accuracy in using two English grammatical structures. *Language Learning*, 64(1), 103-131. doi: <https://doi.org/10.1111/lang.12029>
- Shoufan, A. (2023). Exploring Students' Perceptions of ChatGPT: Thematic Analysis and Follow-Up Survey. *IEEE Access*, 11, 38805-38818. doi : 10.1109/ACCESS.2023.3268224
- Skinner, E. A., Kindermann, T. A. & Furrer, C. J. (2009). A motivational perspective on engagement and disaffection: Conceptualization and assessment of children's behavioral and emotional participation in academic activities in the classroom. *Educational and Psychological Measurement* 69(3), 493-525. <http://dx.doi.org/10.1177/0013164408323233>
- Skinner, E. A., Kindermann, T. A., Connell, J. P., & Wellborn, J. G. (2009). Engagement and disaffection as organizational constructs in the dynamics of motivational development. Dans K. R. Wenzel & A. Wigfield (dir.), *Handbook of motivation at school* (p. 223-245). Routledge/Taylor & Francis Group.
- Slavuj, V., Kovačić, B. & Jugo, I. (2015). *Intelligent Tutoring Systems for Language Learning [Les systèmes de tutorat intelligents pour l'apprentissage des langues]*. Tiré de https://www.researchgate.net/publication/308836139_Intelligent_tutoring_systems_for_language_learning
- Su, Y., Lin, Y. & Lai, C. (2023). Collaborating with ChatGPT in argumentative writing classrooms [Collaborer avec ChatGPT dans les salles de classe d'écriture argumentative]. *Assessing Writing* 57. <https://doi.org/10.1016/j.asw.2023.100752>
- Svalberg, A. M-L. (2009). Engagement with language: Developing a construct. *Language Awareness*, 18(3-4), 242-258. <https://doi.org/10.1080/09658410903197264>
- Talbi, A. & Legros, D. (2018). L'effet de la correction assistée par ordinateur (CAO) sur la construction des compétences orthographiques des apprenants de FLE en Algérie. *Apprentissage des Langues et Systèmes d'Information et de Communication*, 21. <https://doi.org/10.4000/alsic.3255>
- Tang, J. & Rich, C. S. (2017). Automated Writing Evaluation in an EFL Setting: Lessons from China. *JALT CALL*, 13(2), 117-146. <https://doi.org/10.29140/jaltcall.v13n2.215>
- Urbain, T. (2023, 6 février). Après ChatGPT, Google lance son propre robot conversationnel, nommé Bard. *LaPresse*. Tiré de <https://www.lapresse.ca/>
- Van Beuningen, C. G., De Jong, N. H. & Kuiken, F. (2012). Evidence on the effectiveness of comprehensive error correction in second language writing. *Language Learning*, 62(1), 1-41. <https://doi.org/10.1111/j.1467-9922.2011.00674.x>

- Vygotsky, L. S. (1987). *The collected works of LS Vygotsky: Vol. 1, Problems of general psychology*. New York, NY : Plenum Press.
- Wang, P. (2015). Effects of an Automated Writing Evaluation Program: Student Experiences and Perceptions. *Electronic Journal of Foreign Language Teaching*, 12(1), 79-100. Tiré de <https://e-flt.nus.edu.sg/wp-content/uploads/2020/09/v12n12015/wang.pdf>
- Woodworth, J. & Barkaoui, K. (2020). Perspectives on Using Automated Writing Evaluation Systems to Provide Written Corrective Feedback in the ESL Classroom [Les perspectives sur l'utilisation des systèmes d'évaluation de la rédaction automatique pour la rétroaction corrective écrite dans la salle de classe d'ALS]. *TESL Canada Journal*, 39(2), 234-247. <https://doi.org/10.18806/tesl.v37i2.1340>
- Wozney, L., Venkatesh, V. & Abrami, P. C. (2006). Implementing Computer Technologies: Teachers' Perceptions and Practices. *Journal of Technology and Teacher Education*, 14(1), 173-207. Tiré de https://www.researchgate.net/publication/255566914_Implementing_Computer_Technologies_Teachers'_Perceptions_and_Practices
- Wu, H., Wang, W., Wan, Y., Jiao, W. & Lyu, M. R. (2023). *ChatGPT or Grammarly? Evaluating ChatGPT on Grammatical Error Correction Benchmark*. ArXiv. <https://doi.org/10.48550/arXiv.2303.13648>
- Xian, Y., Lampert, C. H., Schiele, B. & Akata, Z. (2020). *Zero-Shot Learning - A Comprehensive Evaluation of the Good, the Bad and the Ugly*. ArXiv. <https://doi.org/10.48550/arXiv.1707.00600>
- Yannakoudakis, H., Andersen, Ø, E., Geranpayeh, A., Briscoe, T. & Nicholls, D. (2018). Developing an Automated Writing Placement System for ESL Learners. *Applied Measurement in Education*, 31(3), 251-267. <https://doi.org/10.1080/08957347.2018.1464447>
- Yilmaz, Y. & Granena, G. (2016). The role of cognitive aptitudes for explicit language learning in the relative effects of explicit and implicit feedback [Le rôle des aptitudes cognitives pour l'apprentissage explicite des langues dans les effets relatifs de la rétroaction explicite et implicite]. *Bilingualism (Cambridge, England)*, 19(1), 147-161. <https://doi.org/10.1017/S136672891400090X>
- Zhang, Z. (2017). Student engagement with computer-generated feedback: a case study. *ELT Journal*, 71(3), 317-328. doi : 10.1093/elt/ccw089
- Zhang, Z., & Hyland, K. (2018). Student engagement with teacher and automated feedback on L2 writing. *Assessing Writing*, 36, 90–102. <https://doi.org/10.1016/j.asw.2018.02.004>

Zouaq, A. (2005). Systèmes Tutoriels Intelligents pour l'Organisation : Gestion de la connaissance et de l'intelligence collective. Tiré de https://site.uottawa.ca/~azouaq/publications/Thesis-Reports/Predoc_Final_Amal_Zouaq.pdf

8 Annexes

Annexe 1 : Plan de cours de FLS 2751

[FLS 2751 A.pdf](#)

Annexe 2 : Consignes du vox pop

Description et évaluation de la tâche 1:



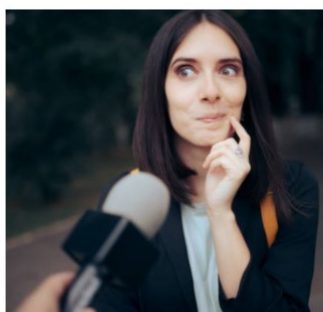
Prendre des risques linguistiques en français

Tâche 1: Le Vox Pop

vox pop n. m.
 vox populi n. f.
 micro-trottoir n. m.

Vox pop est la forme abrégée de *vox populi*, mots latins signifiant « voix du peuple ». Ces deux termes sont invariables : *des vox pop*, *des vox populi*, alors que *micro-trottoir* s'accorde : *des micros-trottoirs*.

Même si *vox pop* est l'abréviation de *vox populi*, qui est de genre féminin, il s'utilise au masculin : *un vox pop*.



Définition

Sondage d'opinion portant généralement sur un sujet d'actualité, effectué dans la rue, pour recueillir les réactions et commentaires spontanés des gens, et dont les résultats sont présentés le plus souvent sous forme de courtes séquences (sonores ou vidéo), par exemple lors d'un journal radiodiffusé ou télévisé.

Exemples:

<https://youtu.be/ot6pxhHJiIE>

<https://youtu.be/jmLDxT58ukI>

<https://youtu.be/Vjrk0FWeQIM>

Comment faire un VOX-POP?



<https://youtu.be/z5BeaGr5SDk>

<https://youtu.be/EoBi3DlpJe8>

Amélioration de la prononciation:

Exercices: [Cliquez ici](#)

Structure de votre Vox-Pop:

Vous pouvez faire votre vox-pop sur le campus ou ailleurs.

A. Sélectionnez un sujet.

B. Développez une question (ou des questions) que vous aimeriez poser à propos de votre sujet. Qu'est-ce que vous désirez prouver dans votre Vox Pop?

C. Gardez la partie technologique simple. Vous pouvez emprunter un micro au pavillon Morisset (sous-sol), salle des multi-médias mais je suggère que vous utilisiez votre téléphone et le fil des écouteurs. Pratiquez une fois ou deux avant de faire les enregistrements.

D. Pour le montage "editing", utilisez l'outil que vous préférez. Par exemple, "clips" dans un iPhone, I Movie, ou vous pouvez utiliser Yuja dans Brightspace.

Voici un court tutoriel pour vous montrer comment l'utiliser

<https://youtu.be/IS0Uw1icT-A>

E. Demandez à un.e ami.e de vous accompagner pour tenir le téléphone pour la vidéo.

8.1 F. Sections de votre Vox Pop:

***Introduction (2-3 minutes)**

Présentez le sujet, le problème, pourquoi avoir choisi ce sujet? que désirez-vous savoir? Opinions? Quelles solutions? Les questions préparées.

***Interviews**

***Vous pouvez insérer des sections où vous parlez entre les interviews.**

***Conclusion (3-4 minutes)**

Analysez les réponses obtenues: Que constatez-vous en écoutant les réponses des gens interviewés, Y a-t-il une tendance? Quel avenir pour ce sujet? Si on vous demandait de faire un autre Vox Pop qui suit celui-ci, qu'aimeriez-vous savoir?.....

***Vous pouvez faire un Vox Pop humoristique 😊**

***Total de minutes: Pas moins de 5 minutes.**

***Date de remise: Entre le 3 et le 15 octobre dans Brightspace (MP4)
Un espace de remise sera créé.**

*** Grille d'évaluation du Vox Pop: Cliquez ici**

Annexe 3 : Questionnaire préintervention

Questionnaire 1

Sélectionnez la réponse avec laquelle vous êtes le plus d'accord. * This form will record your name, please fill your name.

1. Je corrige les textes que j'écris en français.

- Toujours
- Souvent
- Parfois
- Rarement
- Jamais

2. J'utilise les outils traditionnels (les dictionnaires français et les Bescherelle en papier, etc.) pour corriger les textes que j'écris en français.

- Toujours
- Souvent
- Parfois
- Rarement
- Jamais

3. J'utilise les outils numériques pour corriger les textes que j'écris en français.

- Toujours
- Souvent
- Parfois
- Rarement
- Jamais

4. Sélectionnez tous les outils que vous utilisez pour corriger vos textes français. (Cochez toutes les outils que vous utilisez).

- Antidote
- BonPatron
- Google Docs, Microsoft Word ou un autre outil de traitement de texte
- Autre

5. Je suis satisfait des outils de correction linguistique numériques que j'utilise.

- D'accord
- Plutôt d'accord
- Neutre
- Plutôt pas d'accord
- Pas d'accord
- Ne s'applique pas, je n'utilise pas d'outils de correction linguistique numériques.

6. J'ai de la confiance dans les corrections des outils de correction linguistique numériques que j'utilise.

- D'accord
- Plutôt d'accord
- Neutre
- Plutôt pas d'accord
- Pas d'accord
- Ne s'applique pas, je n'utilise pas d'outils de correction linguistique numériques.

7. Je comprends les rétroactions des outils de correction linguistiques numériques que j'utilise.

- D'accord
- Plutôt d'accord
- Neutre
- Plutôt pas d'accord
- Pas d'accord
- Ne s'applique pas, je n'utilise pas d'outils de correction linguistique numériques.

8. J'aimerais pouvoir interagir avec les outils de correction linguistique numériques que j'utilise.

- D'accord
- Plutôt d'accord
- Neutre
- Plutôt pas d'accord
- Pas d'accord
- Ne s'applique pas, je n'utilise pas d'outils de correction linguistique numériques.

9. J'utilise les outils d'intelligence artificielle quand j'écris en français.

- Toujours
- Souvent
- Parfois
- Rarement
- Jamais

10. J'utilise les outils d'intelligence artificielle quand je corrige les textes que j'écris en français.

- Toujours
- Souvent
- Parfois
- Rarement
- Jamais

11. Mon opinion envers l'intelligence artificielle est...

- Positive
- Plutôt positive
- Neutre
- Plutôt négative
- Négative

Annexe 4 : Présentation sur ChatGPT

[ChatGPT pour la correction de l'écrit - FLS2751.pptx](#)

Annexe 5 : Questionnaire post-tâche

Questionnaire 2

Sélectionnez la réponse avec laquelle vous êtes le plus d'accord. * This form will record your name, please fill your name.

1. J'ai aimé utiliser ChatGPT pour corriger mon texte.

- D'accord
- Plutôt d'accord
- Neutre
- Plutôt pas d'accord
- Pas d'accord

2. Utiliser ChatGPT pour la correction a été facile.

- D'accord
- Plutôt d'accord
- Neutre
- Plutôt pas d'accord
- Pas d'accord

3. Utiliser ChatGPT pour la correction a été intuitif.

- D'accord
- Plutôt d'accord
- Neutre
- Plutôt pas d'accord
- Pas d'accord

4. Les explications de ChatGPT ont été utiles pour la correction.

- D'accord
- Plutôt d'accord
- Neutre
- Plutôt pas d'accord
- Pas d'accord

5. Les explications de ChatGPT ont été claires.

- Toujours
- Souvent
- Parfois
- Rarement
- Jamais

6. Les explications de ChatGPT ont été logiques.

- Toujours
- Souvent
- Parfois
- Rarement
- Jamais

7. Les explications de ChatGPT ont été cohérentes.

- Toujours
- Souvent
- Parfois
- Rarement
- Jamais

8. Utiliser ChatGPT pour la correction a été plus facile qu'utiliser d'autres outils de correction linguistique numériques.

- D'accord
- Plutôt d'accord
- Neutre
- Plutôt pas d'accord
- Pas d'accord
- Ne s'applique pas. Je n'utilise pas d'outils de correction linguistique numériques.

9. J'avais de la confiance dans les corrections de ChatGPT.

- D'accord
- Plutôt d'accord
- Neutre
- Plutôt pas d'accord
- Pas d'accord

10. J'ai bénéficié de mes interactions avec ChatGPT lors de la correction.

- D'accord
- Plutôt d'accord
- Neutre
- Plutôt pas d'accord
- Pas d'accord

11. Répondez seulement si vous avez négocié le sens avec ChatGPT. Si vous n'avez pas, continuez sans répondre. En négociant le sens avec ChatGPT, j'ai mieux compris mes erreurs.

- D'accord
- Plutôt d'accord
- Neutre
- Plutôt pas d'accord
- Pas d'accord

12. Comme outil de correction, ChatGPT compte plus d'affordances que de limites.

- D'accord
- Plutôt d'accord
- Neutre
- Plutôt pas d'accord
- Pas d'accord

13. Mon opinion envers l'intelligence artificielle est maintenant....

- Plus positive qu'auparavant
- Inchangée
- Plus négative qu'auparavant

14. Formuler des requêtes pour obtenir de la RCÉ a été facile.

- Toujours
- Souvent
- Parfois
- Rarement
- Jamais

15. Veuillez lister des exemples des requêtes qui ont fourni des rétroactions utiles.

16. Avez-vous continué à utiliser ChatGPT pour l'autoapprentissage après cette expérience?

Oui

Non

17. Répondez seulement si vous avez continué à utiliser ChatGPT pour l'autoapprentissage après cette expérience. Si vous n'avez pas, continuez sans répondre. Comment est-ce que vous avez continué à utiliser ChatGPT pour l'autoapprentissage?

Annexe 6 : Formulaire de consentement

Consentement libre et éclairé

Titre de l'étude : L'utilisation de ChatGPT pour la rétroaction corrective écrite interactive en apprentissage du français langue seconde

Taegan Holmes

Étudiante à la maîtrise

Institut des langues officielles et du bilinguisme

Faculté des arts



Marie-Josée Hamel

Superviseure

Institut des langues officielles et du bilinguisme

Faculté des arts



Invitation à participer :

Je suis invité(e) à participer au projet de recherche mentionné ci-haut mené par Taegan Holmes et supervisé par Dr. Marie-Josée Hamel. Ce projet est pour la thèse de maîtrise de Taegan Holmes.

But de l'étude :

L'étude a pour but de collectionner des données portant sur les pratiques correctives et les attitudes envers l'intelligence artificielle (IA) des étudiants en français langue seconde au niveau postsecondaire, de collectionner un corpus de fils de conversation authentiques entre les étudiants et l'outil d'IA générative, ChatGPT, lors du processus de révision textuelle et de collectionner des données portant sur cette expérience et les perceptions des étudiants concernant les affordances et les limites de l'outil.

Participation :

Ma participation volontaire consistera de répondre à deux questionnaires en ligne et de participer à une activité de révision textuelle interactive avec ChatGPT. Je partagerai le lien du fil de conversation avec la chercheuse principale. Toutes les données seront anonymisées et gardées sur un ordinateur protégé par un mot de passe. Je ne perdrai pas de points dans le cours si je décide de ne pas participer à ce projet.

Chaque questionnaire prendra environ 20 minutes à répondre. L'activité avec ChatGPT prendra environ une heure.

Risques :

La chercheuse principale ne s'attend pas à des risques autres que celles qui pourraient se produire déjà dans la vie quotidienne. Je suis assuré(e) que les données collectionnées seront anonymisées et mon identité ne sera pas divulguée.

Bienfaits :

Ma participation à cette recherche permettra de collectionner des données sur l'utilisation de l'intelligence artificielle générative dans l'enseignement et l'apprentissage du français langue seconde afin d'améliorer les pratiques correctives des enseignants et des apprenants.

Confidentialité et respect de la vie privée :

La chercheuse m'a donné l'assurance qu'elle traitera l'information que je partagerai avec elle de façon strictement confidentielle. Je m'attends à ce que le contenu ne soit utilisé que pour le projet décrit ci-haut. Mon anonymat est préservé par l'anonymisation des données collectionnées et par la suppression de tout document contenant de l'information identifiante. Afin de minimiser les risques de bris de sécurité et pour assurer ma confidentialité, la chercheuse me recommande d'utiliser des mesures de sécurité standard, telles que mettre fin à la session, me déconnecter de mon compte, fermer mon navigateur Internet et verrouiller mon écran ou appareil lorsque je ne les utilise plus.

Conservation des données :

Les données anonymes (réponses aux questionnaires et la transcription du fil de discussion) seront conservées de façon sécuritaire indéfiniment.

Participation volontaire :

Ma participation à cette recherche est volontaire et je suis libre de me retirer en tout temps, de refuser de répondre à toute question à laquelle je ne veux pas répondre sans subir de conséquences négatives. Si je choisis de me retirer de l'étude, le lien de mon fil de conversation sera supprimé et ne sera pas utilisé. En raison des réponses anonymes aux deux sondages, une fois les réponses soumises, je ne pourrai pas retirer ces données de l'étude parce qu'il serait impossible de repérer les données de chaque participant.

Pour tout renseignement additionnel concernant cette étude, je peux communiquer avec la chercheuse ou sa superviseure. Pour tout renseignement sur les aspects éthiques de cette recherche, je peux m'adresser au Bureau d'éthique et d'intégrité de la recherche de Université d'Ottawa au (613) 562-5387 ou ethique@uottawa.ca. La chercheuse recommande de garder une copie du formulaire de consentement.

Acceptation :

En envoyant un courriel à [REDACTED] qui indique « oui, je veux participer à l'étude 'L'utilisation de ChatGPT pour la rétroaction corrective écrite interactive en apprentissage du français langue seconde' » ; je donne mon consentement à participer à cette étude. »

Annexe 7 : Certificat d'approbation éthique

Université d'Ottawa
Bureau d'éthique et d'intégrité de la recherche

17/08/2023
University of Ottawa
Office of Research Ethics and Integrity

CERTIFICAT D'APPROBATION ÉTHIQUE | CERTIFICATE OF ETHICS APPROVAL

Numéro du dossier / Ethics File Number	S-07-23-0440
Titre du projet / Project Title	L'UTILISATION DE CHATGPT POUR LA RÉTROACTION CORRECTIVE ÉCRITE INTERACTIVE EN APPRENTISSAGE DU FRANÇAIS LANGUE SECONDE
Type de projet / Project Type	Thèse de maîtrise / Master's thesis
Statut du projet / Project Status	Approuvé / Approved
Date d'approbation (jj/mm/aaaa) / Approval Date (dd/mm/yyyy)	17/08/2023
Date d'expiration (jj/mm/aaaa) / Expiry Date (dd/mm/yyyy)	16/08/2024

Équipe de recherche / Research Team

Chercheur / Researcher	Affiliation	Role
Taegan HOLMES	Institut des langues officielles et du bilinguisme / Official Languages and Bilingualism Institute	Chercheur Principal / Principal Investigator
Marie-Josée HAMEL	Institut des langues officielles et du bilinguisme / Official Languages and Bilingualism Institute	Superviseur / Supervisor

Conditions spéciales ou commentaires / Special conditions or comments

550, rue Cumberland, pièce 154 550 Cumberland Street, Room 154
Ottawa (Ontario) K1N 6N5 Canada Ottawa, Ontario K1N 6N5 Canada

613-562-5387 • 613-562-5338 • ethique@uOttawa.ca / ethics@uOttawa.ca
www.recherche.uottawa.ca/deontologie | www.recherche.uottawa.ca/ethics

17/08/2023

Université d'Ottawa

Bureau d'éthique et d'intégrité de la recherche

University of Ottawa

Office of Research Ethics and Integrity

Le Comité d'éthique de la recherche (CÉR) de l'Université d'Ottawa, opérant conformément à l'*Énoncé de politique des Trois conseils* (2014) et toutes autres lois et tous règlements applicables, a examiné et approuvé la demande d'éthique du projet de recherche ci-nommé.

L'approbation est valide pour la durée indiquée plus haut et est sujette aux conditions énumérées dans la section intitulée "Conditions Spéciales ou Commentaires". Le formulaire « Renouveau ou Fermeture de Projet » doit être complété quatre semaines avant la date d'échéance indiquée ci-haut afin de demander un renouvellement de cette approbation éthique ou afin de fermer le dossier.

Toutes modifications apportées au projet doivent être approuvées par le CÉR avant leur mise en place, sauf si le participant doit être retiré en raison d'un danger immédiat ou s'il s'agit d'un changement ayant trait à des éléments administratifs ou logistiques du projet. Les chercheurs doivent aviser le CÉR dans les plus brefs délais de tout changement pouvant augmenter le niveau de risque aux participants ou pouvant affecter considérablement le déroulement du projet, rapporter tout événement imprévu ou indésirable et soumettre toute nouvelle information pouvant nuire à la conduite du projet ou à la sécurité des participants.

The University of Ottawa Research Ethics Board, which operates in accordance with the *Tri-Council Policy Statement* (2014) and other applicable laws and regulations, has examined and approved the ethics application for the above-named research project.

Ethics approval is valid for the period indicated above and is subject to the conditions listed in the section entitled "Special Conditions or Comments". The "Renewal/Project Closure" form must be completed four weeks before the above-referenced expiry date to request a renewal of this ethics approval or closure of the file.

Any changes made to the project must be approved by the REB before being implemented, except when necessary to remove participants from immediate endangerment or when the modification(s) only pertain to administrative or logistical components of the project. Investigators must also promptly alert the REB of any changes that increase the risk to participant(s), any changes that considerably affect the conduct of the project, all unanticipated and harmful events that occur, and new information that may negatively affect the conduct of the project or the safety of the participant(s).

Germain ZONGO

Responsable d'éthique en recherche / Protocol Officer

Pour/For Barbara GRAVES Président(e) du/ Chair of the Comité d'éthique de la recherche en sciences sociales et humanités / Social Sciences and Humanities Research Ethics Board

550, rue Cumberland, pièce 154 Ottawa (Ontario) K1N 6N5 Canada

550 Cumberland Street, Room 154
Ottawa, Ontario K1N 6N5 Canada

613-562-5387 • 613-562-5338 • ethique@uOttawa.ca / ethics@uOttawa.ca
www.recherche.uottawa.ca/deontologie | www.recherche.uottawa.ca/ethics