

Holmes, T. (2024). L'utilisation de ChatGPT 3.5 pour la rétroaction corrective écrite interactive en enseignement-apprentissage du français langue seconde : une étude exploratoire. *Actes des Journées de linguistique*, 1, 17-30. <https://doi.org/10.70637/xvc8m613>

© L'auteur-e, 2024. Article en libre accès publié par les *Actes des Journées de linguistique* et diffusé sous licence [Attribution 4.0 International \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/).



Actes des XXXVII^{es} Journées de linguistique | 6 et 7 mars 2024
Université Laval (Québec, Canada)

L'utilisation de ChatGPT 3.5 pour la rétroaction corrective écrite interactive en enseignement-apprentissage du français langue seconde : une étude exploratoire

Taegan Holmes 

Institut des langues officielles et du bilinguisme, Université d'Ottawa
Laboratoire ICAR, Université Lumière Lyon 2

Résumé

Les outils d'intelligence artificielle générative (IAG) deviennent de plus en plus accessibles et conséquemment, il est impératif d'examiner d'un œil critique leur impact potentiel dans le milieu éducatif, ainsi que les manières de les y intégrer pour permettre aux apprenants d'en tirer profit. ChatGPT (OpenAI, 2022) est un outil d'IAG conversationnel avec lequel les utilisateurs peuvent entrer en interaction. La nature interactive de cet outil semble proposer des affordances en enseignement-apprentissage des langues, plus précisément pour la révision textuelle. Cette étude exploratoire porte sur l'utilisation de ChatGPT pour la rétroaction corrective écrite (RCÉ) interactive en français langue seconde. Les participants (n = 22) sont des apprenants de français langue seconde dans un cours de deuxième année universitaire. Ils ont d'abord répondu à un questionnaire portant sur leurs pratiques autocorrectives pour les tâches de rédaction en français. Ensuite, lors d'une intervention ponctuelle, ils ont interagi avec ChatGPT pour solliciter une RCÉ interactive au sein d'échanges synchrones. Enfin, les participants ont répondu à un questionnaire portant sur cette expérience et sur leurs perceptions de l'IA. Une analyse taxonomique (Bilmes, 2009) s'inscrivant dans le cadre de l'analyse qualitative du contenu (Selvi, 2019) a été entreprise pour développer des typologies pour classer les messages dans les fils de discussion recueillis. La qualité des réponses de ChatGPT a aussi été analysée. Les résultats démontrent qu'un grand nombre de différents types de requêtes ont été créés, mais on constate des écarts concernant le degré d'engagement des participants. Cela dit, la grande majorité des participants déclarent avoir tiré profit de cette expérience. Les réponses de ChatGPT étaient majoritairement correctes et appropriées, mais il reste que la qualité de la requête affecte la qualité de la réponse qu'elle sollicite. Les résultats de cette étude semblent démontrer que ChatGPT pourrait être un outil utile pour la RCÉ interactive en apprentissage du français; plus de recherches sont nécessaires.

Mots-clés : ChatGPT, rétroaction corrective écrite, révision textuelle, interaction

Abstract

Generative artificial intelligence (GenAI) tools are becoming increasingly accessible, which makes it imperative to critically examine their potential impact in the educational environment, as well as the ways in which they can be integrated to enable learners to benefit from them. ChatGPT (OpenAI, 2022) is a conversational GenAI tool with which users can interact. The interactive nature of this tool seems to suggest affordances in language teaching and learning, specifically for textual revision. This exploratory study focuses on the use of ChatGPT for interactive written corrective feedback (WCF) in French as a second language. Participants (n=22) were French as a second language learners in a second-year university course. They first answered a questionnaire about their self-corrective practices for French

writing tasks. Then, during a one-off intervention, they interacted with ChatGPT to solicit interactive WCF during synchronous exchanges. Finally, the participants answered a questionnaire about this experience and their perceptions of AI. A taxonomic analysis (Bilmes, 2009) within the framework of a qualitative content analysis (Selvi, 2019) was undertaken to develop typologies to classify the messages in the collected discussion threads. The quality of ChatGPT's responses was also analyzed. The results show that many different types of prompts were created, but there were discrepancies in the degree of participant engagement. That said, the vast majority of participants claim to have benefited from this experience. ChatGPT's responses were largely correct and appropriate, but the quality of the prompt was found to affect the quality of the response it solicits. The results of this study seem to demonstrate that ChatGPT could be a useful tool for interactive WCF in French language learning, however, more research is needed.

Keywords: ChatGPT, written corrective feedback, textual revision, interaction

1. Introduction

De récentes avancées dans le domaine de l'intelligence artificielle (IA) suscitent de l'intérêt dans de nombreux domaines, y compris en enseignement-apprentissage des langues. Aujourd'hui, on s'intéresse surtout aux grands modèles de langage (GML), qui sont des outils d'intelligence artificielle générative (IAG). Les GML sont des outils conversationnels qui permettent aux utilisateurs d'entrer en interaction avec une machine par le biais de requêtes (Office québécois de la langue française, 2023).

Un GML particulièrement populaire est ChatGPT, lancé par l'entreprise *OpenAI* en 2022. ChatGPT a été entraîné à répondre aux instructions des utilisateurs (soumises sous forme de requêtes textuelles) de manière détaillée (OpenAI, 2022). Afin de formuler ses réponses, ChatGPT emploie de nombreuses stratégies, dont un algorithme de prédiction textuelle qui vise à produire la réponse la plus probable à chaque requête.

Les résultats de nombreuses études (voir Fang et al., 2023; Wu et al., 2023) semblent démontrer que ChatGPT propose des affordances en enseignement-apprentissage des langues, surtout en ce qui a trait à la rétroaction corrective écrite (RCÉ). Pour pouvoir s'améliorer en langue seconde, les apprenants ont besoin de recevoir une RCÉ sur leurs productions écrites (Woodworth et Barkaoui, 2020). D'ailleurs, un grand nombre de chercheurs et d'enseignants reconnaissent l'importance de la RCÉ (Bitchener et Ferris, 2011; Hamel et Bibeau, 2021). Bien qu'il existe de nombreux types de RCÉ, il n'y a aucun consensus sur le type le plus efficace. Cela dit, les recherches montrent que la RCÉ synchrone (immédiate) semble être bénéfique (Arroyo et Yilmaz, 2018; Shintani, 2016; Shintani et Aubrey, 2016). La nature conversationnelle de ChatGPT permet justement aux utilisateurs de solliciter une RCÉ qui est non seulement immédiate, mais aussi interactive.

Les recherches récentes sur ChatGPT mettent en lumière plusieurs de ses capacités linguistiques, qui se prêtent particulièrement bien à la RCÉ. D'une part, d'après l'étude de Fang et al. (2023), ChatGPT peut corriger les erreurs grammaticales en anglais, en allemand et en chinois avec un haut degré de précision. Les résultats des études de Coyne et al. (2023) et de Wu et al. (2023) corroborent le fait que ChatGPT est un correcteur linguistique puissant pour la langue anglaise. D'autre part, ChatGPT est en

mesure de fournir des réponses et des rétroactions personnalisées (Holmes, 2024; Su et al., 2023). Enfin, selon les résultats de l'étude de Jiao et al. (2023), ChatGPT possède une bonne capacité de traduction automatique, surtout pour les langues à haut niveau de ressources. Cela dit, à notre connaissance, peu d'études se penchent sur son potentiel didactique pour l'enseignement du français.

Bien que ChatGPT semble proposer de nombreuses affordances, il n'est pas dépourvu de limites. D'abord, en tant que générateur de contenu, ChatGPT tente de produire la réponse la plus probable, ce qui n'est pas toujours la réponse la plus correcte (OpenAI, 2022). La machine peut aussi *fabuler* (Bang et al., 2023; Kaplan-Rakowski et al., 2023), c'est-à-dire produire des informations incorrectes tout en les présentant comme si elles étaient vraies. ChatGPT repose sur des modèles de langage qui utilisent l'apprentissage profond pour prédire le mot suivant dans une séquence, en s'appuyant sur des probabilités calculées à partir d'énormes quantités de données textuelles qui ont servi à son entraînement. Bien que cette approche probabiliste soit efficace pour produire des réponses cohérentes et plausibles, elle peut également générer des informations incorrectes ou inventées. Les résultats des études de Fang et al. (2023) et de Wu et al. (2023) montrent par ailleurs que ChatGPT a tendance à surcorriger.

De concert avec Guichon (2024), nous croyons que, bien que les outils comme ChatGPT aient des limites et qu'il existe des préoccupations autour de leur intégration dans le milieu scolaire, il ne faut pas que les enseignants en interdisent l'accès à leurs étudiants. Au contraire, il faut former les apprenants à les exploiter de manière efficace. Pour ce faire, il est nécessaire de mener des études empiriques pour mieux comprendre les manières utiles et efficaces d'exploiter les nouveaux outils technologiques. Ainsi, cette étude arrive à un moment opportun pour offrir une vue préliminaire et exploratoire sur l'utilisation de ChatGPT pour la RCÉ interactive en apprentissage du français langue seconde (FLS).

2. Méthodes

2.1 Objectifs

Les résultats des études empiriques récentes, ainsi que la nature interactive de ChatGPT (OpenAI, 2022), semblent suggérer que ce dernier pourrait être un outil utile pour fournir une RCÉ interactive aux apprenants de FLS lors d'échanges synchrones. Dans cette optique, les questions de recherche qui suivent ont été formulées :

- 1) Quelles pratiques autocorrectives les participants déclarent-ils utiliser actuellement pour réviser leurs tâches de rédaction?
- 2) Comment les participants se servent-ils de ChatGPT pour obtenir de la RCÉ afin de réviser leur texte écrit?
- 3) Comment ChatGPT répond-il aux requêtes des participants, en termes de précision, pertinence et satisfaction des attentes?

- 4) Les attitudes des participants envers l'IA évoluent-elles après avoir réalisé un exercice impliquant une RCÉ de ChatGPT?

Pour répondre à ces questions de recherche, une étude comportant trois volets a été conçue : un questionnaire prêt-à-remplir portant sur les pratiques de révision textuelle des participants; une tâche de rédaction et de révision interactive avec ChatGPT (version 3.5)¹ et un questionnaire post-tâche portant sur cette expérience. Avant de procéder à la collecte de données, une approbation éthique a été obtenue (numéro du dossier : S-07-23-9440).

2.2 Données

Les participants de cette étude étaient des étudiants de niveau intermédiaire faible (B1) en français inscrits à un cours de FLS de deuxième année à l'Université d'Ottawa (n = 22). La professeure du cours a téléversé un formulaire de consentement sur le site web du cours. Les données collectées et analysées dans la présente étude sont uniquement celles des étudiants ayant donné leur accord de participation. 24 étudiants ont donné leur accord pour participer à cette étude. Toutefois, en raison de difficultés techniques, seules les données de 22 participants étaient utilisables.

Les participants ont d'abord répondu à un questionnaire numérique qui a été partagé sur le site web du cours portant sur leurs pratiques autocorrectives pour les tâches de rédaction en français. Ce questionnaire comprenait des affirmations que les participants ont évaluées avec des échelles de Likert.

Les participants ont pris part à une intervention ponctuelle pendant laquelle ils ont rédigé un court texte qu'ils ont ensuite révisé à l'aide de ChatGPT. Pendant cette intervention, nous leur avons présenté ChatGPT, ainsi qu'un exemple d'interaction avec la machine pour solliciter une RCÉ. Ensuite, les participants ont rédigé leur texte (une introduction à un vox pop de 100 à 150 mots) directement dans ChatGPT. Cette tâche adaptée du cours de FLS auquel ils étaient inscrits n'était pas évaluée. Les participants ont ensuite interagi avec ChatGPT pour réviser leur texte à l'aide de deux requêtes imposées : « 1. Révise ce texte : [introduction] » et « 2. Explique toutes les révisions ». Les étudiants ont été invités à entrer en interaction avec ChatGPT afin de solliciter des précisions, de négocier le sens et la forme et de poser des questions de suivi dans le but de mieux comprendre les rétroactions proposées. Les participants ont disposé d'environ une heure pour rédiger et réviser leur texte à l'aide de ChatGPT. Les liens vers leur fil de discussion avec ChatGPT ont été collectés.

Enfin, les participants ont répondu à un deuxième questionnaire numérique portant sur l'expérience d'utiliser ChatGPT pour la RCÉ, ainsi que sur leurs perceptions de l'IA. Le deuxième questionnaire comprenait des affirmations visant à évaluer la facilité d'utilisation de ChatGPT à l'aide d'échelles de Likert, ainsi qu'une question à réponse ouverte portant sur les types de requêtes qui ont sollicité des rétroactions utiles.

¹ Dans cette étude, à partir de cette note, le terme « ChatGPT » réfère exclusivement à la version 3.5.

2.3 Analyses

Pour analyser les données quantitatives récoltées dans les questionnaires, le décompte des réponses pour chaque question a été effectué. Les fils de discussion ont été recueillis et analysés dans le logiciel NVivo à l'aide d'une analyse taxonomique (Bilmes, 2009) s'inscrivant dans le cadre de l'analyse qualitative du contenu (Selvi, 2019), ce qui a permis l'élaboration de deux typologies pour classer d'abord les différentes requêtes créées par les participants et ensuite les réponses de ChatGPT. Cette méthode d'analyse a été réutilisée pour analyser les réponses à la question ouverte dans le deuxième questionnaire. Les fils de discussion ont aussi été décortiqués pour obtenir des données quantitatives, telles que le nombre de mots par fil de discussion, le nombre de tours de parole, le nombre de requêtes imposées et le nombre de requêtes additionnelles. Enfin, l'alignement entre les requêtes et les réponses, ainsi que la qualité des réponses fournies par ChatGPT, ont été analysés.

3. Résultats

3.1 Analyse du questionnaire préintervention

Le premier questionnaire visait à recueillir des données portant sur les pratiques autocorrectives des participants pour les tâches de rédaction en français. Parmi les 22 participants, 90,1 % ont indiqué toujours ou souvent corriger les textes qu'ils rédigent en français, en utilisant surtout des outils numériques. Seulement 22,7 % des participants ont déclaré toujours ou souvent utiliser les outils traditionnels tels que les dictionnaires papier, tandis que 77,3 % des participants ont déclaré toujours ou souvent utiliser les outils de correction linguistique numériques tels que les correcticiels et les logiciels de traitement de texte. La majorité des participants (86,4 %) ont déclaré être satisfaits de ces outils et avoir confiance dans l'exactitude de leurs corrections. Toutefois, 63,6 % des participants ont déclaré vouloir pouvoir interagir avec les outils qu'ils exploitent. 31,8 % des participants n'utilisent jamais ou rarement les outils d'IA pour corriger les textes qu'ils rédigent en français. Bien que l'utilisation d'outils d'IA ne soit pas une pratique répandue chez les participants, la majorité des participants (54,5 %) ont déclaré avoir des opinions positives ou plutôt positives de l'IA.

Tableau 1. Tableau de synthèse des résultats du premier questionnaire

Affirmation	Toujours	Souvent	Parfois	Rarement	Jamais
Je corrige les textes que j'écris en français.	7	13	1	0	1
J'utilise les outils traditionnels (les dictionnaires français et les Bescherelles en papier, etc.) pour corriger les textes que j'écris en français.	1	4	7	7	3
J'utilise les outils numériques pour corriger les textes que j'écris en français.	8	9	2	0	2
J'utilise les outils d'intelligence artificielle quand je corrige les textes que j'écris en français	2	5	3	3	9

3.2 Analyse des fils de discussion

Avant de procéder à l'analyse du contenu des messages dans les fils de discussion, le corpus de fils de discussion a été décortiqué selon le nombre de mots, le nombre de tours de parole et le nombre de requêtes (imposées et additionnelles). Une synthèse de ces données est présentée dans le Tableau 2 ci-dessous.

Tableau 2. Tableau de synthèse du décompte des fils de discussion

Variable	Moyenne	Écart type	Médiane	Minimum	Maximum
Nombre de mots (total)	787,7	363,5	652,5	399	1656
Nombre de tours de parole	10	4,5	10	4	20
Nombre de requêtes imposées	1,9	0,4	2	1	2
Nombre de nouvelles requêtes	3	2,3	3	0	8

En tout, le corpus comprend 17 331 mots et 220 tours de parole, répartis en 110 requêtes des participants et 110 réponses de ChatGPT. En moyenne, les fils de discussion sont constitués de 787,7 mots ($\pm 363,5$ mots) et 10 tours de parole ($\pm 4,5$ tours). Les participants ont employé en moyenne 1,9 requête imposée ($\pm 0,4$ requête) (P7, P17 et P22 ont omis la deuxième requête imposée) et 3 nouvelles requêtes dans leur fil de discussion ($\pm 2,3$ requêtes). Les participants ont produit 67 nouvelles requêtes. Une première typologie a été développée afin de classer les requêtes des étudiants selon les buts qu'ils visaient à accomplir. Le Tableau 3 ci-dessous résume les catégories de requêtes, leur description ainsi que le nombre de requêtes dans chaque catégorie.

Tableau 3. Tableau de synthèse des requêtes des participants

Catégorie	Description	Nombre de requêtes
Requête imposée	Les deux requêtes imposées : 1) révise ce texte et 2) explique toutes les révisions.	43
Clarification	Les requêtes sollicitant des clarifications et des explications additionnelles portant sur le texte révisé, le texte original, ou concernant les deux à la fois.	42
Socialisation avec la machine	Les requêtes saluant, remerciant ou complimentant ChatGPT.	11
Désaccord	Les requêtes où le participant exprime son désaccord avec ChatGPT.	6
Question autre	Les requêtes posant des questions non liées à la révision textuelle.	3
Révision	Les requêtes sollicitant une révision additionnelle ou portant sur une révision additionnelle.	3
Modification	Les requêtes demandant à ChatGPT de modifier le texte.	2

La catégorie de requêtes la plus fréquente est celle des requêtes imposées (43). Viennent ensuite, avec 42 occurrences, les requêtes de clarification visant à solliciter des précisions et des explications additionnelles. Onze requêtes avaient pour but de socialiser avec la machine (par des salutations, des remerciements ou des compliments), tandis que 6 occurrences reflétaient un désaccord exprimé par les participants envers ChatGPT. Trois requêtes étaient des questions non liées à la révision textuelle (questions autres). Trois autres requêtes sollicitaient ou portaient sur une révision additionnelle, c'est-à-dire une révision autre que la révision initiale sollicitée par la première requête imposée. Enfin, les participants ont demandé à ChatGPT d'effectuer une modification à deux reprises.

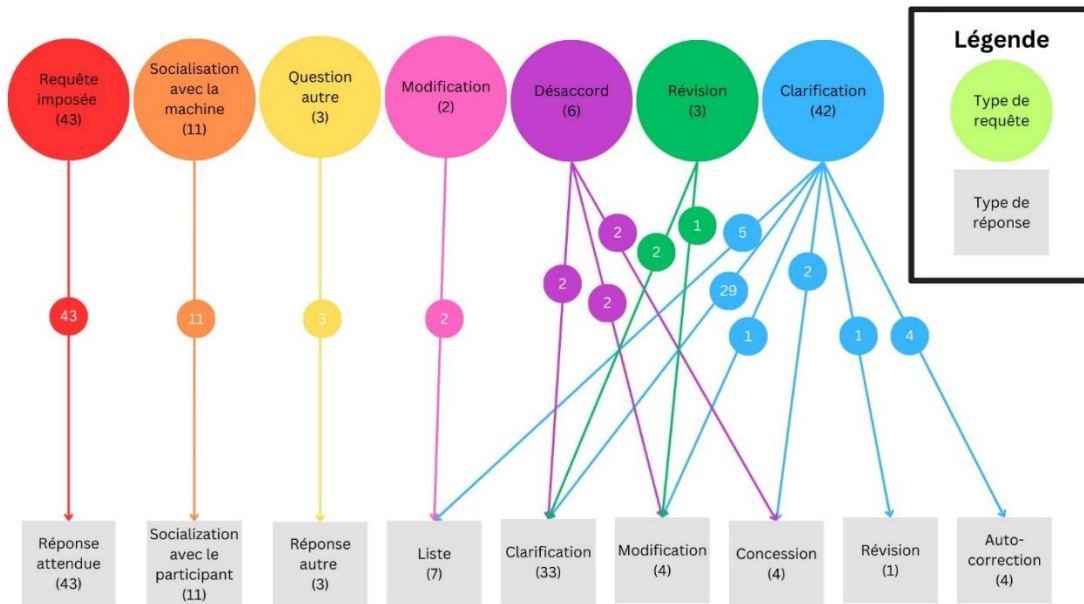
Une deuxième typologie a ensuite été élaborée afin de catégoriser les réponses de ChatGPT en fonction des processus discursifs observés dans les échanges (voir le Tableau 4 ci-dessous).

Tableau 4. Tableau de synthèse des réponses de ChatGPT

Catégorie	Description	Nombre de requêtes
Réponse attendue	Les réponses de ChatGPT répondant aux deux requêtes imposées.	43
Clarification	Les réponses fournissant une précision ou une explication additionnelle portant sur le texte révisé, le texte original, ou concernant les deux à la fois.	33
Socialisation avec le participant	Les réponses de socialisation telles que les salutations, les remerciements, etc.	11
Liste	Les réponses fournissant une liste d'éléments.	7
Concession	Les réponses où ChatGPT concède une erreur ou propose un compromis.	4
Modification	Le texte avec des modifications autres que celles effectuées pendant la révision initiale.	4
Autocorrection	Les réponses où ChatGPT reconnaît sa propre erreur et la corrige.	4
Réponse autre	Les réponses aux questions hors sujet.	3
Révision	Les réponses fournissant le texte révisé une deuxième fois.	1

Les 43 réponses dans la catégorie de réponse attendue répondaient aux deux requêtes imposées avec le texte révisé et les explications des révisions effectuées. 33 réponses de type *clarification* contenaient des précisions ou des explications additionnelles. Onze réponses de ChatGPT étaient des messages de socialisation, comme des salutations et des remerciements. À 7 reprises, ChatGPT a répondu aux participants avec une liste d'éléments, par exemple une liste d'expressions alternatives pour remplacer un extrait dans le texte. ChatGPT a concédé une erreur à quatre reprises (concession) et a répondu avec une version modifiée du texte quatre fois (modification). La machine a reconnu sa propre erreur et l'a corrigée quatre fois (autocorrection). ChatGPT a répondu aux questions de type *autres* trois fois et a partagé le texte révisé une fois (révision).

L'alignement des requêtes et des réponses a été analysé afin de déterminer quels types de requêtes ont entraîné quels types de réponses. La Figure 1 montre une représentation visuelle de l'alignement requête-réponse.

Figure 1. Diagramme de l'alignement requête-réponse

La Figure 1 illustre que certaines catégories de requêtes sont en étroite corrélation avec certaines des catégories de réponses. Par exemple, les requêtes imposées ont toujours sollicité des réponses attendues. Il y a aussi des requêtes dans une même catégorie ayant sollicité des réponses dans différentes catégories. Les 42 requêtes de clarification ont sollicité surtout des réponses de clarification (29), mais également cinq réponses dans la catégorie liste, deux concessions, quatre autocorrections, une modification et une réponse dans la catégorie révision.

Une analyse préliminaire de la qualité des réponses de ChatGPT a aussi été effectuée. Des 110 réponses de ChatGPT, 20 (18,2 %) ont été jugées incorrectes ou inappropriées. La moitié de ces réponses sont soit des cas où ChatGPT prétend avoir effectué une modification qui n'a pas été faite, soit des cas où la machine classe incorrectement une révision effectuée. Quatre des vingt réponses incorrectes étaient des réponses où ChatGPT a fabulé ou a partagé une information incorrecte (à ce sujet, voir Bang et al., 2023 et Kaplan-Rakowski et al., 2023). Enfin, à six reprises, des requêtes mal formulées ou imprécises ont sollicité des réponses inappropriées.

3.3 Analyse du questionnaire post-intervention

Le questionnaire postintervention a porté sur les perceptions des participants à l'égard de l'IA après avoir utilisé ChatGPT pour la RCÉ interactive, ainsi que sur les affordances et les limites perçues de son utilisation pendant le processus de révision textuelle.

La grande majorité des participants (81,2 %) se sont déclarés d'accord ou plutôt d'accord avec l'affirmation selon laquelle ils ont apprécié l'utilisation de ChatGPT pour corriger leur texte. 90,1 % des participants étaient d'accord ou plutôt d'accord

qu'utiliser ChatGPT pour la correction était facile. D'après la majorité des participants, les explications de ChatGPT étaient toujours ou souvent claires (95,5 %), logiques (90,1 %) et cohérentes (95,5 %). 86,4 % des participants ont déclaré être d'accord ou plutôt d'accord d'avoir confiance dans les corrections de ChatGPT. La majorité des participants (54,5 %) n'ont pas changé d'opinion à l'égard de l'IA après avoir utilisé ChatGPT pour la RCÉ interactive, mais parmi ceux ayant changé d'avis, 90 % affichaient désormais une perception de l'IA plus positive qu'auparavant. Enfin, les participants ont indiqué que les requêtes imposées et les requêtes sollicitant une clarification leur étaient les plus utiles.

4. Discussion

Cette étude exploratoire s'est intéressée à l'utilisation de ChatGPT 3.5 pour la rétroaction corrective écrite interactive dans le contexte de l'enseignement-apprentissage du français langue seconde. Elle s'articulait autour de quatre questions de recherche :

- 1) Quelles pratiques autocorrectives les participants déclarent-ils utiliser actuellement pour réviser leurs tâches de rédaction?
- 2) Comment les participants se servent-ils de ChatGPT pour obtenir de la RCÉ afin de réviser leur texte écrit?
- 3) Comment ChatGPT répond-il aux requêtes des participants, en termes de précision, pertinence et satisfaction des attentes?
- 4) Les attitudes des participants envers l'IA évoluent-elles après avoir réalisé un exercice impliquant une RCÉ de ChatGPT?

Nous abordons chacune de ces questions dans les sous-sections qui suivent.

4.1 QR1 : Quelles pratiques autocorrectives les participants déclarent-ils utiliser actuellement pour réviser leurs tâches de rédaction?

L'analyse des réponses au premier questionnaire a permis de constater que la majorité des participants corrigent les textes qu'ils rédigent en français et privilégient, pour ce faire, les outils de correction numériques. Ils ont déclaré être majoritairement satisfaits des outils qu'ils exploitent et ont exprimé un haut degré de confiance dans les corrections offertes par ces outils.

4.2 QR2 : Comment les participants se servent-ils de ChatGPT pour obtenir de la RCÉ afin de réviser leur texte écrit?

L'analyse des fils de discussion a relevé que certains participants ont été plus engagés que d'autres pendant la révision textuelle avec ChatGPT. P5, P9 et P14 n'ont envoyé que les requêtes imposées et ne sont pas entrés en interaction avec la machine pour solliciter des précisions, pour poser des questions de suivi ou pour négocier le sens, ainsi démontrant un faible degré d'engagement avec la tâche linguistique (Egbert,

2020). En revanche, certains participants comme P18, P19 et P20 ont interagi avec ChatGPT pour discuter des rétroactions sollicitées, ce qui semble démontrer plus d'engagement de leur part (Egbert, 2020). D'après les résultats de l'étude de Koltovskaia (2020), les apprenants de langue seconde ont souvent des pratiques autocorrectives uniques. Il n'est donc pas surprenant qu'il existe un si grand écart entre les degrés d'engagement des participants pendant leur échange avec ChatGPT.

4.3 QR3 : Comment ChatGPT répond-il aux requêtes des participants, en termes de précision, pertinence et satisfaction des attentes?

Pour sa part, ChatGPT a bien effectué la révision des textes soumis lors de la révision initiale. Cela corrobore les résultats de Coyne et al. (2023), de Fang et al. (2023) et de Wu et al. (2023) qui montrent que la machine peut effectuer la correction des erreurs grammaticales dans une langue à haut niveau de ressources (p. ex. l'anglais et le français) avec un haut degré de précision. De concert avec Fang et al. (2023) et Loem et al. (2023), les résultats de cette étude démontrent que la qualité de la requête a un impact significatif sur la qualité de la réponse qu'elle sollicite.

L'analyse de l'alignement requête-réponse a révélé que ChatGPT semble adopter une approche de cas par cas pour formuler ses réponses, en tenant compte des informations incluses dans les différentes requêtes pour leur répondre de manière pertinente et précise tout en évitant l'emploi de réponses préformatées qui pourraient manquer de pertinence. Ceci corrobore le fait que ChatGPT est en mesure de fournir des réponses et des rétroactions personnalisées (Su et al., 2023). Cependant, les requêtes imprécises ou mal formulées peuvent solliciter des réponses inappropriées ou incorrectes (Su et al., 2023), ce qui souligne l'importance de bien formuler les requêtes.

4.4 QR4 : Les attitudes des participants envers l'IA évoluent-elles après avoir réalisé un exercice impliquant une RCÉ de ChatGPT?

D'après les résultats au deuxième questionnaire, la majorité des participants n'ont pas changé d'opinion envers l'IA après avoir utilisé ChatGPT pour la RCÉ interactive. Toutefois, de ceux ayant changé d'avis, 90 % ont désormais une opinion plus positive qu'auparavant. Les participants ont déclaré avoir majoritairement aimé cette expérience et ont déclaré avoir tiré profit de leur échange avec ChatGPT. Cela dit, quelques participants ont rencontré des difficultés en ce qui concerne l'utilité, la clarté, la logique et la cohérence des réponses de ChatGPT.

5. Conclusion

Les résultats de cette étude exploratoire suggèrent que ChatGPT pourrait être un outil utile pour la RCÉ interactive en apprentissage du FLS. D'une part, ChatGPT est un correcteur linguistique puissant, capable de réviser des textes en français avec un haut degré de précision et de fournir des rétroactions personnalisées. Sa nature conversationnelle permet aux utilisateurs d'entrer en interaction pendant la révision

textuelle pour solliciter des précisions et pour poser des questions de suivi. D'autre part, les participants à cette étude ont majoritairement tiré profit de leur échange avec la machine et ont déclaré avoir aimé cette expérience. Il est cependant essentiel de s'assurer que les étudiants (tout comme les enseignants) comprennent les affordances et les limites de ChatGPT pour la révision textuelle. Il est aussi important que les étudiants comprennent la différence entre utiliser ChatGPT comme outil pour la révision textuelle et l'utiliser pour effectuer son travail sans s'y impliquer personnellement. Les résultats de cette étude sont prometteurs, mais davantage de recherches sont nécessaires pour arriver à des conclusions plus sûres.

De futures recherches devraient inclure un plus grand échantillon, composé de participants de différents niveaux de compétence linguistique. Il serait aussi pertinent de proposer une plus grande variété de tâches aux participants, incluant par exemple la rédaction d'un texte argumentatif ou d'une histoire fictive dans diverses conditions de réalisation (tâche notée ou non notée, contraintes de temps, etc.). Il serait également pertinent de mener des études sur les nouvelles versions de ChatGPT au fur et à mesure de leur parution.

Nombre de mots (incluant les notes de bas de page) : 3815

6. Matériel connexe

Les évaluations éditoriales et l'évaluation anonyme de l'article sont stockées sur Figshare :

<https://doi.org/10.6084/m9.figshare.27001282.v2>

7. Remerciements

Je tiens à remercier ma superviseuse de thèse, la Professeure Marie-Josée Hamel, ainsi que les évaluateurs du présent article.

8. Financement et conflits d'intérêts

L'auteure déclare qu'elle n'a aucun conflit d'intérêt réel, potentiel ou perçu.

9. Références

- Arroyo, D. C. et Yilmaz, Y. (2018). An open for replication study: The role of feedback timing in synchronous computer-mediated communication. *Language Learning*, 68(4), 942-972. <https://doi.org/10.1111/lang.12300>
- Bang, Y., Cahyawijaya, S., Lee, N., Dai, W., Su, D., Wilie, B., Lovenia, H., Ji, Z., Yu, T., Chung, W., Do, Q. V., Xu, Y. et Fung, P. (2023). A multitask, multilingual, multimodal evaluation of ChatGPT on reasoning, hallucination, and

- interactivity. Dans J. C. Park, Y. Arase, B. Hu, W. Lu, D. Wijaya, A. Purwarianti, A. A. Krisnadhi (dir.), *Proceedings of the 13th International Joint Conference on Natural Language Processing and the 3rd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics (Volume 1: Long Papers)* (p. 675-718). Association for Computational Linguistics.
<https://doi.org/10.18653/v1/2023.ijcnlp-main.45>
- Bilmes, J. (2009). Taxonomies are for talking: Reanalyzing a Sacks classic. *Journal of Pragmatics*, 41(8), 1600-1610. <https://doi.org/10.1016/j.pragma.2008.10.008>
- Bitchener, J. et Ferris, D. (2011). *Written corrective feedback in second language acquisition and writing*. Routledge.
- Coyne, S., Sakaguchi, K., Galvan-Sosa, D., Zock, M. et Inui, K. (2023). Analyzing the performance of GPT-3.5 and GPT-4 in grammatical error correction. *ArXiv*. Prépublication. <https://doi.org/10.48550/arXiv.2303.14342>
- Egbert, J. (2020). Language task engagement (theory). Dans J. Egbert et M. F. Roe (dir.), *Theoretical models for teaching and research*. Pressbooks.
<https://opentext.wsu.edu/theoreticalmodelsforteachingandresearch/chapter/language-task-engagement-theory/>
- Fang, T., Yang, S., Lan, K., Wong, D. F., Hu, J., Chao, L. S. et Zhang, Y. (2023). Is ChatGPT a highly fluent grammatical error correction system? A comprehensive evaluation. *ArXiv*. Prépublication. <https://doi.org/10.48550/arXiv.2304.01746>
- Guichon, N. (2024). *Sur les traces de Richard Kern: Acknowledging the pivotal role of technologies in language education*. *The Modern Language Journal*, 108(2), 563-566. <https://doi.org/10.1111/modl.12931>
- Hamel, M. J. et 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>
- Holmes, T. (2024). ChatGPT pour la rétroaction corrective écrite interactive en apprentissage du français langue seconde [thèse de maîtrise, Université d'Ottawa]. Thèses uOttawa. <https://doi.org/10.20381/ruor-30407>
- Jiao, W., Wang, W., Huang, J.-t., Wang, X., Shi, S. et Tu, Z. (2023). Is ChatGPT a good translator? Yes with GPT-4 as the engine. *ArXiv*. Prépublication.
<https://doi.org/10.48550/arXiv.2301.08745>
- Kaplan-Rakowski, R., Grotewold, K., Hartwick, P. et Papin, K. (2023). Generative AI and teachers' perspectives on its implementation in education. *Journal of Interactive Learning Research*, 34(2), 313-338.
<https://www.learntechlib.org/p/222363/>
- Koltovskaia, S. (2020). Student engagement with automated written corrective feedback (AWCF) provided by *Grammarly*: A multiple case study. *Assessing Writing*, 44, article 100450. <https://doi.org/10.1016/j.asw.2020.100450>

- Loem, M., Kaneko, M., Takase, S. et Okazaki, N. (2023). Exploring effectiveness of GPT-3 in grammatical error correction: A study on performance and controllability in prompt-based methods. Dans E. Kochmar, J. Burstein, A. Horbach, R. Laarmann-Quante, N. Madnani, A. Tack, V. Yaneva, Z. Yuan et T. Zesch (dir.), *Proceedings of the 18th Workshop on Innovative Use of NLP for Building Educational Applications (BEA 2023)* (p. 205-219). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2023.bea-1.18>
- Office québécois de la langue française. (2023). Grand modèle de langage. Dans *Grand dictionnaire terminologique*. <https://vitrinelinguistique.oqlf.gouv.qc.ca/fiche-gdt/fiche/26570880/grand-modele-de-langage>
- OpenAI. (2022, 30 novembre). *Introducing ChatGPT*. <https://openai.com/blog/chatgpt>
- Selvi, A. F. (2019). Qualitative content analysis. Dans J. McKinley et H. Rose (dir.), *The Routledge handbook of research methods in applied linguistics* (p. 440-452). Routledge.
- 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. <https://doi.org/10.1080/09588221.2014.993400>
- Shintani, N. et Aubrey, S. (2016). The effectiveness of synchronous and asynchronous written corrective feedback on grammatical accuracy in a computer-mediated environment. *The Modern Language Journal*, 100(1), 296-319. <https://doi.org/10.1111/modl.12317>
- Su, Y., Lin, Y. et Lai, C. (2023). Collaborating with ChatGPT in argumentative writing classrooms. *Assessing Writing*, 57, article 100752. <https://doi.org/10.1016/j.asw.2023.100752>
- Woodworth, J. et Barkaoui, K. (2020). Perspectives on using automated writing evaluation systems to provide written corrective feedback in the ESL classroom. *TESL Canada Journal*, 37(2), 234-247. <https://doi.org/10.18806/tesl.v37i2.1340>
- Wu, H., Wang, W., Wan, Y., Jiao, W. et Lyu, M. (2023). ChatGPT or Grammarly? Evaluating ChatGPT on grammatical error correction benchmark. *ArXiv*. Prépublication. <https://doi.org/10.48550/arXiv.2303.13648>