

Faculté de philosophie, arts et lettres

LA SIMPLIFICATION DES TEXTES LITTÉRAIRES EN FLE À L'ÈRE DE L'IA : DE L'ÉDITION EN FRANÇAIS FACILE AUX VERSIONS SIMPLIFIÉES GÉNÉRÉES PAR CHATGPT. UNE ÉTUDE COMPARATIVE À PROPOS DE L'ŒUVRE D'ÉMILE ZOLA, *AU BONHEUR DES DAMES*

Auteure : Hanh Tam Nguyen

Promoteurs : Séverine De Croix

Van Toan Nguyen

Année académique 2025-2026

Master [120] en langues et lettres françaises et romanes,
orientation français langue étrangère, à finalité spécialisée

Déclaration sur l'honneur, mémoire de Master de la Faculté de philosophie, arts et lettres, UCLouvain

Je déclare qu'il s'agit d'un travail original et personnel, que toutes les sources référencées ont été indiquées dans leur totalité et ce, quelle que soit leur provenance, et que l'utilisation des outils d'intelligence artificielle est conforme aux consignes générales d'utilisation publiées par l'UCLouvain et disponible à cette adresse :

[Consignes-ChatGPT-version-etudiante.pdf](#)

Je suis conscient·e que le fait de ne pas citer une source, de ne pas la citer clairement et complètement, de ne pas annoncer dans une partie spécifique en début du mémoire l'utilisation que j'ai faite des outils d'intelligence artificielle ou de les utiliser de manière contraire à ce qui est prévu dans la note institutionnelle constituent des irrégularités graves au regard du *Règlement des études et des examens de l'Université catholique de Louvain* (Chapitre 4, section 7, article 107 à 114) au sein de l'Université. J'ai notamment pris connaissance des risques de sanctions académiques et disciplinaires encourues.

Au vu de ce qui précède, je déclare sur l'honneur ne pas avoir commis de plagiat ou toute autre forme de fraude et de n'avoir pas utilisé les outils d'intelligences artificielles en dehors de ce qui est accepté à l'UCLouvain.

Nom, Prénom : Nguyen, Hanh Tam

Date : 31 juillet 2026

Signature de l'étudiant·e :



Nguyen Hanh Tam

REMERCIEMENT

La réalisation de ce mémoire n'aurait pas été possible sans le soutien et les encouragements de nombreuses personnes, auxquelles je souhaite exprimer ma sincère gratitude.

Mes premiers remerciements s'adressent à Madame la Professeure Séverine De Croix, ma promotrice, pour la qualité de son accompagnement tout au long de ce travail. Je lui suis profondément reconnaissante pour sa disponibilité, la pertinence de ses conseils et l'attention qu'elle a toujours portée à mes recherches. Au-delà de son expertise scientifique, sa confiance, sa patience et ses encouragements ont été une véritable source de motivation. Dans les moments où j'avais le sentiment de ne plus avancer, elle a toujours trouvé les mots pour me redonner confiance et m'encourager à poursuivre ce projet avec sérénité.

Je souhaite également remercier Monsieur Nguyễn Văn Toàn, mon directeur de mémoire à l'Université de Hanoï. Je lui adresse toute ma reconnaissance pour sa disponibilité et son soutien constant. Chaque fois que je rencontrais une difficulté, il prenait le temps de répondre à mes questions et de m'orienter avec beaucoup de patience. Ses conseils m'ont permis de progresser à chaque étape de cette recherche. Je remercie chaleureusement l'ensemble des enseignants de la Faculté de Langue françaises de l'Université de Hanoï pour les connaissances qu'ils m'ont transmises tout au long de mon parcours universitaire. Je souhaite également exprimer ma gratitude aux enseignants de l'UCLouvain, dont les cours et les échanges ont enrichi ma réflexion et ouvert de nouvelles perspectives sur la recherche en didactique du français.

J'adresse enfin mes sincères remerciements aux vingt-huit enseignants de FLE qui ont accepté de participer à cette étude. Le temps qu'ils ont consacré à l'évaluation des textes ainsi que la richesse de leurs commentaires ont constitué une contribution essentielle à cette recherche.

À toutes ces personnes, j'adresse ma plus sincère reconnaissance.

TABLE DES MATIÈRES

LISTE DES ABRÉVIATIONS.....	iv
LISTE DES TABLEAUX.....	v
LISTE DES FIGURES.....	vi
RÉSUMÉ.....	vii
INTRODUCTION.....	1
Cadre contextuel.....	1
Problématique.....	3
États des recherches.....	5
Objectifs de recherche.....	9
Questions de recherche.....	10
Hypothèses de recherche.....	11
CHAPITRE 1 : CADRE THÉORIQUE.....	12
1.1. La littérature en FLE et les enjeux de la réception.....	12
1.1.1. Évolution des approches didactiques : du patrimoine au sujet lecteur.....	12
1.1.2. Apports des théories de la réception.....	16
1.1.3. Médiation pédagogique et plaisir de lire.....	20
1.2. Le défi de l'accessibilité : théories de la simplification des textes.....	24
1.2.1. L'input compréhensible et la médiation linguistique.....	24
1.2.2. Définition et niveaux de simplification.....	28
1.2.3. Les opérations de simplification.....	30
1.3. Simplifier Zola : enjeux de littérarité et pertes stylistiques.....	33
1.3.1. Le projet naturaliste de Zola.....	33
1.3.2. Le style zolien comme système.....	34
1.3.3. Que perd-on en simplifiant Zola ?.....	37
1.4. L'IA comme agent de simplification.....	40
1.4.1. Simplification éditoriale et simplification automatique.....	40
1.4.2. Fonctionnement des modèles de langage.....	43
1.4.3. Le prompt comme outil de contrôle.....	46
CHAPITRE 2 : MÉTHODOLOGIE.....	50

2.1. Introduction du chapitre	50
2.1.1. Justification d'une approche hybride et contrastive	50
2.1.2. Présentation générale du dispositif méthodologique	52
2.2. Constitution et description du corpus.....	54
2.2.1. Texte source : <i>Au Bonheur des Dames</i> d'Émile Zola	54
2.2.2. Critères de sélection des quatorze extraits	56
2.2.3. La version éditoriale : l'adaptation de CLE Internationale (Niveau A2).....	62
2.2.4. La version artificielle : protocole de génération de l'IA.....	64
2.3. Grille d'analyse multidimensionnelle.....	66
2.3.1. Analyse linguistique formelle	70
2.3.2. Typologie des procédés de simplification	72
2.3.3. Tension entre fidélité sémantique et littérarité.....	75
2.3.4. Évaluation de la lisibilité objective.....	77
2.4. Étude empirique : le regard de l'expert FLE	82
2.4.1. Objectifs de l'évaluation par perception humaine	82
2.4.2. Échantillonnage : profil des participants	85
2.4.3. Enquête à l'aveugle et codifications des extraits	86
2.4.4. Élaboration du questionnaire et des échelles d'appréciation	88
2.5. Traitement des données	91
CHAPITRE 3 : ANALYSE DES RÉSULTATS ET DISCUSSION	94
3.1. Analyse quantitative de la lisibilité objective	94
3.1.1. Comparaison des indicateurs structurels (ASL et ASW)	94
3.1.2. Comparaison des scores de lisibilité (Flesch et Kandel-Moles).....	95
3.1.3. Analyse du profil lexical et corrélation avec les indices de lisibilité.....	97
3.2. Analyse des procédés de simplification	100
3.2.1. Suppression et découpage syntaxique.....	100
3.2.2. Substitution lexicale et reformulation.....	102
3.2.3. Explicitation.....	104
3.3. Analyse de la préservation des caractéristiques littéraires	106
3.3.1. Préservation des invariants narratifs	106

3.3.2. Préservation des effets sensoriels et descriptifs	108
3.3.3. Le compromis entre l'esthétique narrative zolienne et l'accessibilité du texte.....	110
3.4. Analyse des résultats expérimentaux : le regard des enseignants du FLE	113
3.4.1. Profil des participants	113
3.4.2. Évaluation globale des extraits simplifiés	116
3.4.3. Comparaison des perceptions des enseignants selon le type de simplification.....	118
3.4.4. Influence du profil des enseignants sur les évaluations.....	121
3.4.5. Analyse qualitative des commentaires ouverts	124
3.5. Discussion et recommandation	127
3.5.1. Synthèse et mise en perspective des résultats	127
3.5.2. Forces et limites de la simplification par l'IA	130
3.5.3. Influence de l'ingénierie du prompt sur les résultats obtenus	132
3.5.4. Vers un modèle de simplification hybride adapté au niveau A2.....	133
3.5.5. Limites de la recherche	135
CONCLUSION	138
BIBLIOGRAPHIE.....	142

LISTE DES ABRÉVIATIONS

Abréviation	Signification
ASL	<i>Average Sentence Length</i>
ASW	<i>Average Syllables per Word</i>
CECRL	Cadre européen commun de référence pour les langues
FLE	Français Langue Étrangère
FRE	<i>Flesch Reading Ease</i>
IA	Intelligence artificielle
LLM	<i>Large Language Model</i>
S–V–O	Sujet – Verbe – Objet
VE	Version éditoriale
VIA	Version générée par l'intelligence artificielle
VO	Version originale
ZPD	Zone proximale de développement

LISTE DES TABLEAUX

Tableau 1. <i>Inventaire et contenu narratif des quatorze extraits du corpus</i>	57
Tableau 2. <i>Grille d'analyse multidimensionnelles des textes</i>	68
Tableau 3. <i>Interprétation des scores de Flesch adaptée d'après Flesch (1948)</i>	79
Tableau 4. <i>Dimension expérimentale : critères d'évaluation du questionnaire</i>	90
Tableau 5. <i>Échelle de Likert pour l'évaluation des extraits simplifiés</i>	90
Tableau 6. <i>ASL et ASW moyens des 14 extraits du corpus</i>	94
Tableau 7. <i>Scores de lisibilité des versions du corpus</i>	96
Tableau 8. <i>Moyennes des catégories lexicales selon les trois versions du corpus</i> ..	97
Tableau 9. <i>Taux de suppression textuelle observés dans les versions simplifiées</i> ..	102
Tableau 10. <i>Variations des perceptions sensorielles entre VE et VIA</i>	109
Tableau 11. <i>Évaluation des extraits selon les quatre critères</i>	117
Tableau 12. <i>Classement des extraits selon les scores globaux des enseignants</i>	118
Tableau 13. <i>Comparaison des moyennes générales entre la VE et la VIA</i>	119
Tableau 14. <i>Scores moyens des versions selon le statut linguistique des enseignants</i>	121
Tableau 15. <i>Scores moyens des versions selon l'ancienneté des enseignants</i>	122
Tableau 16. <i>Scores moyens des versions selon le contexte d'enseignement</i>	123
Tableau 17. <i>Proposition d'un modèle hybride de simplification pour le niveau A2</i>	134

LISTE DES FIGURES

Figure 1. <i>Articulation du cadre théorique de la recherche</i>	49
Figure 2. <i>Schéma de la méthodologie de recherche</i>	52
Figure 3. <i>Organisation et critères de sélection du sous-corpus</i>	62
Figure 4. <i>Répartition des réponses obtenues à notre questionnaire (avril 2026)</i>	114
Figure 5. <i>Enquête, suite (pratique du FLE)</i>	115
Figure 6. <i>Enquête, suite (ancienneté)</i>	115
Figure 7. <i>Enquête, suite (contexte)</i>	116

RÉSUMÉ

Cette recherche porte sur la simplification des textes littéraires en Français Langue Étrangère (FLE) à l'ère de l'intelligence artificielle générative. Elle propose une analyse comparative entre une version éditoriale issue de la collection *Français Facile* (CLE International) et une version générée par ChatGPT, à partir d'extraits du roman *Au Bonheur des Dames* d'Émile Zola destinés à des apprenants de niveau A2. La méthodologie adoptée repose sur une approche triangulée combinant trois volets complémentaires : une analyse quantitative de la lisibilité et du profil lexical, une analyse qualitative des procédés de simplification, ainsi qu'une évaluation à l'aveugle menée auprès de 28 enseignants de FLE.

Les résultats mettent en évidence deux logiques de simplification distinctes. Les versions éditoriales privilégient une forte accessibilité linguistique grâce à une réduction importante des phrases et des informations, au risque d'atténuer certaines dimensions de la littérarité. À l'inverse, ChatGPT tend à mieux préserver la progression narrative, les descriptions et l'atmosphère de l'œuvre tout en modernisant le lexique. L'évaluation révèle également que, si les enseignants jugent les versions éditoriales plus accessibles sur le plan linguistique, ils accordent aux versions générées par l'IA un potentiel pédagogique plus important et une plus grande capacité à favoriser les interactions en classe.

Cette recherche montre ainsi que l'intelligence artificielle ne se substitue pas à l'expertise de l'enseignant, mais constitue un outil de médiation susceptible d'accompagner la simplification des textes littéraires tout en contribuant à préserver leur valeur pédagogique et littéraire.

Mots-clés : Français Langue Étrangère (FLE) ; simplification textuelle ; intelligence artificielle générative ; ChatGPT ; didactique de la littérature ; Émile Zola.

ABSTRACT

This study examines the simplification of literary texts in the context of Teaching French as a Foreign Language (FFL) in the era of generative artificial intelligence. It presents a comparative analysis between an editorial adaptation from the *Français Facile* collection (CLE International) and a version generated by ChatGPT, based on selected excerpts from Émile Zola's novel *Au Bonheur des Dames* intended for A2-level learners.

The research adopts a triangulated methodology combining three complementary approaches: a quantitative analysis of readability (Flesch and Kandel-Moles readability indices) and lexical profile (Lextutor), a qualitative analysis of text simplification strategies (deletion, substitution, reformulation, explicitation, and segmentation), and a blind evaluation conducted with 28 teachers of French as a Foreign Language.

The findings reveal two distinct approaches to text simplification. Editorial versions prioritize linguistic accessibility through substantial sentence reduction and information deletion, at the risk of diminishing certain aspects of literary quality. By contrast, ChatGPT tends to preserve the narrative progression, descriptive passages, and atmosphere of the original work while modernizing the vocabulary. The evaluation further shows that, although teachers consider the editorial versions more linguistically accessible, they attribute greater pedagogical value and a stronger potential for promoting classroom interaction to the AI-generated versions.

This study therefore suggests that artificial intelligence does not replace teachers' expertise but rather serves as a mediation tool that can support the simplification of literary texts while helping preserve their pedagogical and literary value.

Keywords: French as a Foreign Language (FFL); text simplification; generative artificial intelligence; ChatGPT; literature didactics; Émile Zola.

INTRODUCTION

Cadre contextuel

L'enseignement du FLE a beaucoup évolué au cours des dernières années. Aujourd'hui, apprendre une langue ne consiste plus uniquement à mémoriser des règles de grammaire ou du vocabulaire. Les apprenants doivent aussi être capables de comprendre des documents authentiques, de communiquer dans différentes situations et de développer progressivement leur autonomie.

Cette évolution amène les enseignants à réfléchir davantage aux supports qu'ils utilisent en classe. Le choix des supports dépend en grande partie du niveau des apprenants. Un texte adapté à des étudiants de niveau B2 peut s'avérer trop difficile pour une classe de niveau A2. La littérature pose souvent cette question. Elle donne accès à la langue, mais aussi à la culture, à une époque et à une certaine vision du monde. C'est pourquoi de nombreux enseignants cherchent à l'intégrer dans leurs cours, même avec des apprenants débutants.

Pourtant, les œuvres patrimoniales restent souvent difficiles d'accès pour les apprenants de FLE. Les difficultés ne viennent pas seulement des mots inconnus. Les phrases sont parfois très longues, les descriptions nombreuses et les références historiques ou culturelles demandent des connaissances que les étudiants ne possèdent pas toujours. Cette situation apparaît clairement dans plusieurs universités vietnamiennes où les étudiants travaillent principalement sur des extraits proposés par leurs enseignants. Les œuvres intégrales sont rarement étudiées en classe. Lorsque nous enseignons le français, nous avons souvent constaté que certains étudiants perdaient rapidement le fil d'un texte de Hugo ou de Maupassant. Ils s'arrêtaient sur des mots comme « boutique », « commis » ou « bourgeoisie », puis finissaient par se concentrer uniquement sur la traduction du vocabulaire. À ce moment-là, ils ne lisaient plus vraiment l'histoire.

Pour faciliter cet accès, les maisons d'édition proposent depuis longtemps des versions adaptées dans des collections de *Français Facile*. Ces adaptations réduisent les principales difficultés linguistiques tout en conservant le fil de l'histoire. Elles

sont aujourd'hui largement utilisées en FLE, car elles rendent les œuvres plus accessibles à des apprenants qui ne pourraient pas encore lire la version originale. L'apparition des outils d'intelligence artificielle générative a ouvert une autre possibilité. Il est désormais possible de demander à un système comme ChatGPT de produire une version simplifiée d'un texte en précisant le niveau des apprenants et quelques consignes de réécriture. Pour un enseignant qui prépare plusieurs cours par semaine, ce gain de temps est particulièrement intéressant. Il devient possible de produire rapidement un texte adapté ou même des activités pédagogiques. Simplifier un texte littéraire ne consiste pas seulement à remplacer des mots difficiles par des mots plus simples. Chaque modification peut transformer le style, le rythme ou encore l'atmosphère du récit. Certains passages perdent leur force parce que les descriptions sont raccourcies. D'autres deviennent plus faciles à comprendre, mais aussi plus neutres. On peut alors se demander jusqu'où il est possible de simplifier une œuvre sans lui faire perdre son caractère littéraire.

Ces interrogations sont particulièrement importantes dans le contexte vietnamien. Les enseignants disposent souvent de peu de temps pour créer leurs propres supports. Les adaptations éditoriales ne sont pas toujours disponibles dans les bibliothèques, et leur prix reste parfois élevé pour les établissements ou les étudiants. Dans ces conditions, l'intelligence artificielle apparaît comme une solution accessible et rapide. Pourtant, son utilisation ne peut pas reposer uniquement sur un critère de rapidité. Il est également nécessaire de s'intéresser à la qualité des textes produits.

Les adaptations proposées par les éditeurs et celles générées par l'intelligence artificielle poursuivent le même objectif : rendre une œuvre plus accessible aux apprenants de FLE. En revanche, elles ne reposent pas sur la même démarche. Les premières sont le résultat d'un véritable travail éditorial, tandis que les secondes sont produites automatiquement à partir d'instructions données à un modèle d'intelligence artificielle.

Problématique

Si l'intelligence artificielle est aujourd'hui capable de simplifier rapidement un texte littéraire, une question demeure ouverte : cette simplification est-elle comparable à celle réalisée par les éditions pédagogiques ?

Par ailleurs, il ne faut pas oublier que la simplification des textes littéraires n'est pas une pratique nouvelle née avec l'IA. Depuis longtemps, le FLE s'appuie sur des collections de *Français Facile* ou de lectures graduées, dans lesquelles des éditeurs adaptent des œuvres littéraires pour les rendre accessibles aux apprenants. L'arrivée de l'intelligence artificielle générative modifie aujourd'hui cette pratique. Alors que la simplification reposait jusqu'à présent sur le travail des éditeurs, ChatGPT peut désormais produire une version simplifiée en quelques secondes. Cette évolution conduit à s'interroger sur les différences entre ces deux façons de simplifier un même texte.

En effet, derrière cette efficacité technologique se cache une question plus large. Simplifier un texte littéraire ne consiste pas seulement à réduire sa difficulté linguistique. Chaque choix de reformulation modifie aussi l'équilibre entre la compréhension, le maintien de certaines caractéristiques du texte et les possibilités de son exploitation en classe. Une adaptation peut devenir plus accessible sans répondre pour autant aux objectifs pédagogiques recherchés. À l'inverse, une simplification moins poussée permet de préserver davantage d'éléments du texte d'origine, mais elle peut rendre la lecture plus exigeante pour un public de niveau A2. Cette difficulté apparaît clairement dans *Au Bonheur des Dames* d'Émile Zola. Les descriptions y sont nombreuses et le vocabulaire naturaliste occupe une place importante. Ces éléments participent pleinement à l'écriture du roman, mais ils constituent aussi une source de difficulté pour des apprenants de FLE. Les éditions en *Français Facile* et les versions générées par ChatGPT répondent-elles à cette difficulté de la même manière ? Les choix de simplification opérés par l'une et l'autre conduisent-ils à des textes comparables du point de vue de leur lisibilité, de leurs caractéristiques linguistiques et de leur utilisation possible en classe de FLE ?

Cette réflexion dépasse finalement la seule question du style. Un texte simplifié peut être linguistiquement très clair sans être perçu comme un bon support pédagogique. À l'inverse, un texte un peu plus dense peut offrir davantage de possibilités d'échanges ou d'activités en classe. La qualité d'une simplification ne dépend donc pas d'un seul critère, mais d'un équilibre entre plusieurs dimensions.

Ce danger ne touche pas seulement les auteurs mais aussi les lecteurs, surtout les étudiants. Un lecteur ne lit pas seulement pour recevoir une information ; il lit pour vivre une expérience. Barthes (1973), dans *Le Plaisir du texte*, explique très bien que le plaisir littéraire vient de ce mélange particulier entre l'effort, la surprise et le jeu du langage. Mais lorsque le texte devient trop simple, trop direct, ce plaisir disparaît. C'est un peu comme un plat sans saveur : il nourrit, mais il ne donne aucune émotion. Un texte simplifié de manière trop mécanique — que ce soit par un éditeur maladroit ou par une IA — produit exactement cet effet : il transmet l'histoire mais supprime tout ce qui pourrait donner envie de lire davantage.

C'est dans cet espace de manque que se situe notre recherche. Les études actuelles sur l'IA sont nombreuses mais très générales : on y parle de grammaire, de vocabulaire, de feedback, comme le montrent par exemple les travaux de Kohnke et al. (2023). Très peu de travaux confrontent directement les textes simplifiés générés par l'IA aux éditions en *Français Facile* en analysant simultanément leurs caractéristiques linguistiques, leurs choix de simplification et leur potentiel didactique. On sait que l'IA peut simplifier, mais on ne sait pas encore si elle peut rivaliser avec le travail d'adaptation éditoriale sur le plan du style et de l'émotion.

C'est pourquoi la question cruciale de ce mémoire n'est plus de savoir comment rédiger un prompt, mais d'évaluer le résultat : **Dans quelle mesure la simplification opérée par ChatGPT se distingue-t-elle de l'édition traditionnelle en *Français Facile* dans l'adaptation d'*Au Bonheur des Dames* destinée à des apprenants de FLE de niveau A2, tant sur le plan linguistique que pédagogique ?** C'est précisément à partir de cette tension entre deux modèles — l'expert humain et l'algorithme — que notre travail prend forme, pour mieux comprendre ce que chaque

mode de simplification apporte, ce qu'il laisse de côté et dans quelles conditions ces textes peuvent constituer des ressources pertinentes pour l'enseignement du FLE.

États des recherches

La simplification des textes littéraires en FLE ne constitue pas un sujet nouveau. Bien avant l'apparition de l'intelligence artificielle générative, cette question occupait déjà une place importante dans les recherches consacrées à l'enseignement des langues : comment permettre à des apprenants dont la maîtrise du français demeure limitée d'accéder à des œuvres littéraires sans attendre un niveau avancé de compétence ?

La simplification d'un texte ne se limite pas au choix des mots. Modifier une œuvre implique aussi de revoir la syntaxe, l'organisation des informations, les références culturelles ou encore le rythme de lecture. Autrement dit, simplifier ne consiste pas seulement à remplacer des mots difficiles par des termes plus fréquents. À partir d'un corpus de fables françaises, Safira et al. (2019) montrent que ce travail relève d'une véritable réécriture. Chaque modification suppose de trouver un équilibre entre une meilleure lisibilité et le respect du texte d'origine. Cette réflexion est particulièrement importante dans l'enseignement de la littérature. Rendre une œuvre plus accessible ne signifie pas seulement en diminuer les difficultés linguistiques. Comme le souligne Kayembe (2019), la littérature conserve une fonction formatrice, même pour des lecteurs débutants. Elle participe au développement de l'autonomie, de la sensibilité et de l'esprit critique. Une adaptation trop poussée risquerait donc d'effacer une partie de ce qui fait l'intérêt de l'œuvre. Les collections en *Français Facile* illustrent bien cette recherche d'équilibre. Selon Lovrečki (2023), elles occupent une position intermédiaire entre le manuel de langue et la lecture intégrale. Les apprenants peuvent ainsi découvrir des œuvres patrimoniales dans une version adaptée, puis passer progressivement aux textes originaux.

Malgré la diversité de leurs approches, ces travaux convergent sur un point essentiel. La simplification n'est pas pensée comme un appauvrissement de la littérature. Elle constitue une médiation. Son objectif n'est pas d'effacer les caractéristiques du texte littéraire, mais de rendre possible une première rencontre avec l'œuvre en tenant compte des compétences réelles des lecteurs. C'est dans cette conception que s'inscrit

le développement des collections en *Français Facile*, aujourd'hui largement intégrées aux pratiques d'enseignement du FLE.

Nous constatons pourtant que les études disponibles portent presque exclusivement sur les adaptations produites par les maisons d'édition, et que les conséquences d'une simplification réalisée par un modèle d'intelligence artificielle restent beaucoup moins documentées. Or, l'apparition d'outils comme ChatGPT modifie profondément les possibilités de production et d'adaptation des textes. Si la simplification éditoriale est aujourd'hui bien connue, l'usage de l'IA générative dans ce domaine ouvre encore de nombreuses questions auxquelles la littérature apporte peu de réponses.

L'essor de l'intelligence artificielle générative a profondément renouvelé les recherches consacrées à la production automatique de textes. Pourtant, les travaux récents ne permettent pas encore de dégager un consensus clair sur les capacités réelles de ces modèles. Les recherches récentes aboutissent à des conclusions parfois contradictoires, qui invitent à considérer les performances de l'IA avec davantage de nuance. Ainsi, Rinck (2025), dans son article intitulé *Des textes de l'IA générative pour former à l'enseignement de la production écrite*, insiste sur l'intérêt pédagogique des textes générés par l'IA lorsqu'ils sont utilisés comme objets d'analyse. L'analyse de ces productions amène les futurs enseignants à porter leur attention sur des aspects plus larges que la simple correction linguistique. Ils s'intéressent aussi à l'organisation du texte, à sa cohérence et à son fonctionnement discursif. L'étude de Leroy et al. (2024) aborde la question sous un autre angle, à partir d'un corpus de textes d'information médicale simplifiés. Les auteurs montrent qu'un expert humain évalue plus favorablement le style des textes générés par ChatGPT que celui des versions simplifiées manuellement. En revanche, la préservation du contenu obtient des résultats moins satisfaisants. Une rédaction plus fluide ne suffit donc pas à garantir le maintien des informations essentielles : un texte peut gagner en qualité stylistique tout en s'éloignant davantage du contenu de la version d'origine.

La question est abordée sous un autre angle par Yang (2025). Alors que plusieurs études portent principalement sur la qualité des textes produits par l'IA, cet auteur s'intéresse davantage à ses capacités créatives. Il remet notamment en question l'idée selon laquelle les systèmes d'IA se limiteraient à reproduire des modèles existants. Selon lui, les modèles les plus récents peuvent faire émerger un « potentiel créatif caché », que les méthodes d'évaluation habituelles ne permettent pas toujours de mettre en évidence. L'IA n'est donc plus envisagée uniquement comme un outil de reproduction, mais aussi comme un système capable, dans certaines situations, de produire des formes originales de créativité. Les conclusions de ces études ne sont donc pas directement comparables, puisqu'elles reposent sur des objectifs et des critères d'évaluation différents. Cette diversité montre surtout que les performances de l'IA varient selon l'aspect observé : la qualité stylistique, la fidélité au texte source ou encore la créativité ne sont pas nécessairement évaluées de la même manière. Cette observation est particulièrement importante dans le cadre de notre recherche, qui s'intéresse précisément à la simplification d'un texte littéraire plutôt qu'à la génération libre d'un nouveau texte.

Ces travaux montrent que les productions de l'IA peuvent être jugées convaincantes par des lecteurs humains. Ils apportent toutefois encore peu d'éléments concernant la simplification des œuvres littéraires. Générer un texte nouveau et adapter une œuvre littéraire ne relèvent pas des mêmes opérations. La distinction est importante. Dans le second cas, la difficulté ne consiste pas seulement à produire un texte lisible, mais à transformer un texte existant tout en négociant un équilibre entre accessibilité, fidélité et préservation de ses caractéristiques littéraires. C'est précisément ce niveau d'analyse qui demeure encore peu documenté.

Cette question conduit naturellement à revenir sur les recherches consacrées à la simplification textuelle. En effet, comparer plusieurs versions d'un même texte suppose de disposer d'outils permettant de décrire précisément les transformations réalisées ; les travaux développés dans ce domaine offrent justement un cadre méthodologique permettant d'observer ces modifications de façon plus systématique.

La simplification des textes ne constitue pas un sujet apparu avec l'intelligence artificielle générative. Depuis plusieurs décennies déjà, elle occupe une place importante dans les recherches en didactique des langues et alimente des réflexions aussi bien théoriques que pédagogiques.

Parmi les contributions les plus utiles à notre étude, Gala et al. (2018) proposent un cadre d'analyse particulièrement structuré. Les auteurs définissent la simplification de manière opératoire et distinguent quatre niveaux d'intervention : le lexique, la morphologie, la syntaxe et le discours. Au-delà de cette approche linguistique, la question de la simplification fait depuis longtemps l'objet d'un débat didactique. Widdowson (1978) la considère avant tout comme un moyen de faciliter l'accès aux textes et d'accompagner les apprentissages. Dans le champs de la didactique du français langue première, Rouxel (2007), pour sa part, rappelle qu'une simplification excessive peut modifier l'expérience de lecture et conduire à l'effacement de certaines caractéristiques propres aux œuvres littéraires. Cette tension reste encore très présente dans les recherches actuelles.

Malgré le nombre croissant de publications consacrées à l'intelligence artificielle générative, les recherches qui mettent directement en regard ces outils et les adaptations éditoriales restent encore limitées. Cette observation est particulièrement frappante dans le domaine de la littérature en FLE. Au cours de notre revue de littérature, nous avons trouvé de nombreuses études portant soit sur les adaptations en *Français Facile*, soit sur les performances des modèles d'IA, mais très peu qui confrontent les deux à partir d'un même corpus littéraire. Il reste donc difficile de savoir si les transformations opérées suivent des logiques comparables ou si elles conduisent à des résultats sensiblement différents. Cette lacune justifie le choix de notre objet d'étude, qui consiste à comparer une adaptation éditoriale et une version produite par ChatGPT à partir des mêmes extraits d'*Au Bonheur des Dames*, avant d'examiner la manière dont ces deux versions sont perçues par des enseignants de FLE.

Nous voudrions donc proposer de comparer, à partir d'un même corpus littéraire, une version éditoriale et une version générée par ChatGPT afin d'identifier les procédés

de simplification privilégiés par chacune d'elles et d'examiner la manière dont ces choix sont évalués par des enseignants de FLE.

Objectifs de recherche

L'objectif général de cette recherche est d'évaluer et de comparer la qualité des simplifications produites par ChatGPT et par des éditions pédagogiques traditionnelles à partir d'extraits d'*Au Bonheur des Dames* d'Émile Zola, en prenant en compte leur lisibilité, leur potentiel didactique ainsi que leur capacité à préserver les principales dimensions littéraires de l'œuvre.

Pour atteindre cet objectif général, la recherche se décline en trois objectifs opérationnels :

- Analyser les différences linguistiques et discursives entre les simplifications produites par ChatGPT et celles proposées par les éditions pédagogiques : cet objectif consiste à mettre en évidence les principales convergences et divergences entre les deux types de simplification. L'analyse porte notamment sur les procédés de simplification employés, les transformations lexicales, syntaxiques et discursives, ainsi que sur leur influence sur la lisibilité et sur la préservation des principales caractéristiques du texte littéraire.
- Évaluer le potentiel didactique des textes simplifiés auprès d'enseignants de FLE : cet objectif vise à recueillir le regard d'enseignants de FLE sur les différentes versions simplifiées afin d'apprécier leur pertinence pour un usage en classe. L'évaluation prend en compte plusieurs dimensions, notamment la clarté des textes, leur adéquation à un public de niveau A2, leur potentiel d'exploitation pédagogique ainsi que leur capacité à favoriser les interactions et les activités d'apprentissage.
- Identifier les apports et les limites de ChatGPT dans la simplification des textes littéraires par rapport aux adaptations éditoriales : à partir des analyses linguistiques et des évaluations réalisées par les enseignants, cet objectif cherche à déterminer dans quelle mesure les textes générés par ChatGPT peuvent constituer une alternative crédible aux adaptations éditoriales. Il s'agit également de mettre en évidence les points forts et les limites de chaque

approche afin de mieux cerner leurs possibilités d'utilisation dans l'enseignement de la littérature en FLE.

Questions de recherche

Sur la base de la problématique posée et des objectifs définis, cette étude s'articule autour d'une interrogation centrale :

Dans quelle mesure la simplification opérée par ChatGPT se distingue-t-elle de l'édition traditionnelle en *Français Facile* dans l'adaptation d'*Au Bonheur des Dames* destinée à des apprenants de FLE de niveau A2, tant sur le plan linguistique que pédagogique ?

Pour répondre à cette question centrale, notre réflexion se décompose en trois questions spécifiques :

- Quelles sont les principales divergences morphosyntaxiques et stylistiques observables entre les versions simplifiées par l'édition pédagogique traditionnelle et celles produites par l'IA ?
- Selon les enseignants, quelle version présente les caractéristiques les plus proches d'un texte littéraire exploitable en classe de FLE ?
- L'IA tend-elle à gommer systématiquement certains éléments littéraires au nom de la clarté, là où l'éditeur humain choisit délibérément de les conserver pour des raisons esthétiques ?

Hypothèses de recherche

Pour répondre à ces questions, nous avons formulé les hypothèses de recherche suivantes :

- La première hypothèse (H1) suppose que les simplifications générées par ChatGPT favoriseront une meilleure accessibilité linguistique grâce à une plus grande régularité lexicale et syntaxique. En contrepartie, elles pourraient conduire à une certaine standardisation du style et à une atténuation de certaines caractéristiques littéraires présentes dans le texte source de Zola, en comparaison avec les adaptations éditoriales.
- La deuxième hypothèse (H2) suppose que les stratégies de simplification adoptées par ChatGPT diffèrent de celles des adaptations éditoriales. Alors que ces dernières tendraient à préserver davantage certains éléments lexicaux et culturels du texte source, l'IA privilégierait plus fréquemment des reformulations visant à maximiser la lisibilité, au risque d'atténuer certaines spécificités de l'œuvre originale.

CHAPITRE 1 : CADRE THÉORIQUE

Ce premier chapitre pose les bases théoriques de la recherche et met en évidence les tensions qui traversent la didactique de la littérature en Français Langue Étrangère (FLE), notamment entre la transmission des œuvres, les différentes formes de simplification des textes et l'usage croissant de l'intelligence artificielle (IA). À partir de là, une question centrale se dégage : dans quelle mesure l'IA peut-elle contribuer à la simplification des textes de Zola pour des apprenants de niveau A2, en comparaison avec les pratiques éditoriales traditionnelles, sans altérer l'expérience de lecture littéraire ? Ce chapitre s'appuie sur plusieurs approches, allant des théories de la réception aux travaux sur la simplification textuelle et aux recherches plus récentes sur les usages de l'IA.

1.1. La littérature en FLE et les enjeux de la réception

1.1.1. Évolution des approches didactiques : du patrimoine au sujet lecteur

Dans le champ de la didactique du FLE, la place accordée à la littérature ne va pas de soi. Son intégration en classe demande d'abord de clarifier ce qu'on appelle un texte littéraire, puis de réfléchir aux conditions dans lesquelles ce type de texte peut être proposé à des apprenants non natifs. Cette réflexion nous conduit à interroger la nature même du support utilisé.

Définir l'œuvre littéraire comme un « document authentique », Cuq (2003, p. 29) implique déjà une certaine forme d'exclusion de l'apprenant. En désignant « tout message élaboré par des francophones pour des francophones » dans une visée de « communication réelle », Cuq (2003) met en lumière, peut-être sans le vouloir, une tension importante en didactique : c'est justement cette authenticité qui constitue sa première barrière. Pour un lecteur allophone, cela ne représente pas vraiment une liberté de lecture. Il se retrouve plutôt face à un texte parfois opaque, dont il peut comprendre la syntaxe, mais dont les codes de lecture et d'interprétation lui échappent en partie.

Malgré cette extériorité du lecteur, la littérature reste fortement présente dans les dispositifs de formation, dans la mesure où celle-ci est investie d'une mission de

transmission culturelle et de légitimation des savoirs scolaires. L'œuvre n'est donc pas seulement considérée comme un matériau linguistique, mais comme un « véhicule de valeurs esthétiques et morales » (Jenny, 2010, p. 182) censé garantir la « cohésion d'une communauté » (David, 2010, p. 79). Cependant, cette valeur n'est ni naturelle ni inhérente au texte, car elle résulte d'un processus de légitimation institutionnelle. Comme l'analysent Ronveaux et Schneuwly (2018, p. 6), l'école est une instance qui « invente sans cesse la littérature » en lui attribuant une « réputation littéraire ». À travers ce travail de transformation, le texte brut devient un objet d'enseignement soumis à un processus de « disciplinarisation » (p. 11), c'est-à-dire une mise en forme scolaire qui prédétermine la rencontre entre l'élève et l'œuvre.

En didactique du FLE, une question se pose assez naturellement : comment faire pour que la littérature reste un patrimoine à transmettre, sans pour autant enfermer le sens ni rendre la lecture trop passive, et en même temps permettre aux apprenants de vraiment s'appropriier les textes ? C'est justement ce décalage entre la littérature comme héritage et sa réception en classe qui est au cœur de notre réflexion.

C'est précisément cette tension que Dufays (2020, p. 3) tente de dépasser en proposant un déplacement de centre de gravité. Pour lui, il s'agit de « (re)donner du sens et du goût à l'enseignement [...] en le centrant non plus sur une lecture modèle incarnée par le maître, mais sur l'activité effective de l'élève ». Ce changement est davantage paradigmatique que simplement méthodologique, car il ne se limite pas à une évolution des pratiques, mais redéfinit surtout la place du lecteur dans l'activité de lecture. En réalité, la lecture n'est pas un savoir que l'on transmet de manière stable ; elle se construit dans l'engagement du sujet, et cet engagement est souvent instable.

Ce déplacement redéfinit également le rôle de l'enseignant. Alors qu'il était auparavant le garant d'une interprétation légitime dans une logique plutôt descendante, il devient désormais un médiateur de l'expérience de lecture. L'apprenant, de son côté, n'est plus seulement destinataire du sens. Il entre progressivement dans une position d'acteur, capable de produire du sens à partir du

texte. Une autre conception de la lecture se dessine, moins centrée sur la conformité interprétative que sur l'activité réelle du lecteur.

Au-delà d'un simple changement dans la manière de lire, c'est l'ensemble des pratiques en didactique des langues qui se trouve transformé. En inscrivant dans la perspective actionnelle du CECRL, cité par Puren (2006, p. 7), l'apprenant quitte sa robe de destinataire pour endosser celle, plus exigeante, d'« acteur social ». Lire un texte littéraire ne se limite plus à un exercice scolaire ou purement analytique. C'est désormais une activité située, qui mobilise différentes compétences et s'inscrit dans une dynamique d'action. La lecture peut ainsi être pensée comme une pratique à part entière, et non plus uniquement comme un savoir théorique.

Toutefois, la dimension subjective de la lecture y reste peu prise en compte. Comme la lecture ne peut être réduite à un simple décodage linguistique, les travaux de Annie Rouxel et Gérard Langlade nous permettent de repenser le rôle actif du lecteur. En introduisant la notion de « sujet lecteur », ils déplacent encore le regard. Si Rouxel (2007, p. 69) pose le cadre en affirmant que « l'investissement subjectif du lecteur est une nécessité fonctionnelle de la lecture littéraire ; c'est le lecteur qui achève le texte et lui imprime sa forme singulière », cette idée ne naît cependant pas dans le champ didactique lui-même. Elle s'inscrit dans le prolongement des théories de la réception d'Eco (1965), d'Iser (1985) et de Jauss (1990), que Rouxel et Langlade réinvestissent ensuite dans une réflexion centrée sur l'enseignement de la littérature. Langlade (2014, p. 48) semble néanmoins vouloir pousser cette intuition plus loin. Pour lui, la réception n'est pas qu'un achèvement, c'est un « processus interactionnel » où le lecteur ne se contente pas de remplir les blancs du texte, mais réécrit l'œuvre à partir de sa propre mémoire. Ce dialogue entre les deux auteurs suggère que la littérature ne tient pas seulement à la complexité des phrases. Elle se joue également dans la liberté du lecteur. La simplification ne doit donc pas être vue a priori comme un appauvrissement, mais peut être considérée comme un moyen de permettre à l'apprenant A2 de passer d'un simple déchiffrement à une véritable activité de lecture créatrice.

Cependant, ces modèles du « sujet lecteur », élaborés dans des contextes scolaires natifs, supposent souvent une maîtrise linguistique et culturelle qui n'est pas évidente en FLE. Dans le cas d'apprenants allophones de niveau A2, il faut donc adapter ce regard didactique. Si, avec Rouxel et Langlade, on considère que le sens se construit dans la rencontre avec le lecteur, cette rencontre devient souvent difficile pour ces apprenants, en raison de leurs ressources linguistiques limitées. Les difficultés de vocabulaire et de grammaire agissent comme un frein qui empêche le lecteur de s'engager réellement dans le texte, limitant ainsi son activité interprétative. La simplification n'apparaît donc pas comme une solution unique, mais un outil parmi d'autres permettant de réguler cette difficulté. Dès lors, l'enjeu n'est plus seulement de comprendre comment le sens se construit, mais de créer les conditions qui rendent cette naissance possible.

C'est à ce niveau que l'entrée précoce dans la lecture littéraire devient un défi nécessaire. Réserver l'accès aux œuvres classiques aux seuls apprenants de niveau avancé reviendrait à priver les débutants d'une source majeure de motivation, mais aussi d'une première expérience sensible de la littérature. Pour que ce choix pédagogique soit réellement viable, il est alors nécessaire de repenser la manière dont les textes sont proposés et adaptés. L'idée qui guide notre réflexion est que la complexité de la langue ne résume pas, à elle seule, la richesse d'une œuvre. En allégeant certaines contraintes lexicales ou syntaxiques, on permet à l'esprit de l'apprenant de se libérer du poids de la traduction mot à mot pour porter davantage son attention à la progression du récit, aux images ou aux émotions produites par le texte. La simplification n'est pourtant pas sans conséquences. En cherchant à faciliter la lecture, on peut aussi modifier certains effets stylistiques, expliciter des éléments laissés implicites ou réduire des zones d'ambiguïté qui participent à la lecture de l'œuvre. Pour cette raison, la simplification ne peut pas être considérée comme une simple intervention technique. Elle implique toujours des choix qui influencent, au moins en partie, les caractéristiques du texte original. Elle correspond plutôt à une recherche d'équilibre — parfois instable — entre accessibilité linguistique et préservation de ce qui fait la singularité sensible et littéraire du texte.

Cette réflexion ne peut cependant s'arrêter à la seule question de l'accessibilité linguistique. Elle conduit également à interroger une tension plus profonde, entre les droits du texte et ceux du lecteur. Cette opposition rejoint la « tension » identifiée par Dufays (2017) entre deux pôles de réception. D'un côté, la « distanciation » privilégie les « droits du texte » et les vertus d'une « lecture analytique » ; de l'autre, la « participation » valorise une « lecture subjective » et existentielle, centrée sur les « droits du lecteur » et les « questions vives de l'humain » (Dufays, 2017, p. 7).

Pour un public de niveau A2, cette distinction prend une importance particulière. Un texte original trop complexe pousse souvent l'apprenant vers une lecture distante et difficile, centrée sur la forme, sans réelle possibilité d'accéder aux « questions vives ». La simplification joue ainsi un rôle de médiation entre l'œuvre et le lecteur. Elle cherche à réduire les difficultés linguistiques sans faire disparaître tout ce qui nourrit la réflexion ou l'interprétation. L'objectif n'est donc pas seulement de faciliter la compréhension, mais aussi de permettre aux apprenants d'entrer plus facilement dans l'univers de l'auteur et de construire progressivement leur propre lecture de l'œuvre. Cette réflexion, en posant les bases d'une compréhension nuancée de la simplification, ne se limite pas à ouvrir une perspective pour la suite de cette recherche. Elle constitue surtout le cadre conceptuel à partir duquel seront envisagées les différentes formes d'adaptation des textes de Zola pour des apprenants de niveau A2.

1.1.2. Apports des théories de la réception

Il est difficile aujourd'hui d'évoquer le « sujet lecteur » sans faire référence à l'esthétique de la réception. Mais il ne s'agit pas seulement d'une théorie parmi d'autres : c'est plutôt une manière différente de voir la lecture. Le texte n'est plus considéré comme un objet qui contient un sens fixe, que l'on transmet simplement au lecteur. Au contraire, le sens se construit dans la rencontre entre le texte et le lecteur, à travers la manière dont ce dernier investit, comprend et interprète le texte.

Chez Hans Robert Jauss (1990), cette rencontre n'est jamais neutre, ni totalement libre d'ailleurs. En fait, elle est déjà orientée avant même que la lecture ne commence, et elle ne se produit pas dans un vide théorique. Toute œuvre s'inscrit dans ce que

Jauss appelle un « horizon d'attente », défini comme le « système de références objectivement formulable » issu de la connaissance des genres, des thèmes et de la distinction entre langage poétique et réalité quotidienne (Jauss, 1990, p. 49). Plus précisément, lire revient à activer un ensemble d'attentes : certaines sont conscientes, mais beaucoup d'autres restent implicites. Le lecteur anticipe, reconnaît, compare, parfois même sans s'en rendre compte.

Cela signifie aussi que la lecture littéraire ne relève pas d'une compétence purement spontanée. Comme le rappelle Melançon (1997, p. 6), en s'appuyant notamment sur les travaux de Gagnon (1983) et de Vandendorpe (1989), savoir lire ne signifie pas encore savoir lire une œuvre littéraire. Celle-ci suppose un savoir-faire spécifique, impliquant le déchiffrement des indices, des fonctions narratives et des inférences nécessaires à la construction du sens. Autrement dit, si le lecteur n'est jamais "neutre", il n'est pas non plus immédiatement compétent : la lecture littéraire s'apprend.

Dans le contexte du FLE, cette question se pose de manière encore plus concrète. Les apprenants ne partagent pas toujours les références linguistiques et culturelles auxquelles le texte fait appel. Certains éléments, qui paraissent évidents pour un lecteur natif, demandent davantage d'efforts d'interprétation. Un écart peut alors apparaître entre les connaissances présupposées par le texte et celles que les apprenants sont réellement en mesure de mobiliser au moment de la lecture. Du coup, la difficulté ne vient pas seulement du texte lui-même. Elle se situe plutôt dans cet écart entre les implicites du texte et les ressources interprétatives du lecteur.

On peut alors considérer que la simplification ne consiste pas simplement à rendre un texte plus facile. D'une certaine manière, elle agit aussi sur ce que Hans Robert Jauss appelle l'« horizon d'attente » (Jauss, 1990, p. 49), c'est-à-dire l'ensemble des références culturelles, linguistiques et littéraires que le lecteur mobilise au moment de la lecture. La simplification ne transforme pas complètement cet horizon, mais elle peut en réduire partiellement l'écart avec celui du texte, afin de rendre la rencontre avec le texte plus accessible. C'est dans cette perspective que Jauss évoque ce qu'il appelle la « fusion des deux horizons » (Jauss, 1990, p. 259), entendue comme un rapprochement entre l'univers du texte et celui du lecteur. Dans la pratique, cette

médiation passe par des choix très concrets. La simplification consiste souvent à remplacer quelques mots difficiles par un vocabulaire plus courant. Les changements paraissent parfois modestes, mais ils peuvent suffire à rendre le texte plus accessible. Le même principe s'observe pour la syntaxe. Une tournure comme « il est constaté que... » est plus facilement comprise lorsqu'elle devient « on constate que... ». À elles seules, ces modifications ne transforment pas le texte, mais elles en facilitent la lecture.

Ce geste peut donner l'impression d'être assez technique, presque automatique parfois. Mais en y regardant de plus près, il renvoie quand même à une certaine conception de la lecture et de l'enseignement. Réduire la distance est nécessaire, surtout dans un cadre pédagogique, mais sans effacer ce qui fait la singularité du texte. Et c'est là qu'apparaît une première tension, pas toujours facile à gérer : trop rapprocher le texte du lecteur revient à risquer d'aplatir le style et de neutraliser les nuances ; ne pas intervenir du tout, en revanche revient à laisser le lecteur à distance, dans une situation où le sens devient difficilement accessible.

Ce cadre posé par Jauss ne suffit cependant pas à décrire ce qui se passe réellement dans l'acte de lecture. Il ouvre une voie, mais il laisse aussi beaucoup de zones floues. C'est à ce niveau que l'analyse de Wolfgang Iser entre en jeu : elle complique la réflexion, mais elle l'enrichit également. Le regard change alors : il ne s'agit plus seulement de savoir avec quoi le lecteur arrive, mais de comprendre ce que le texte lui demande de faire concrètement.

Dans *L'acte de lecture*, Iser (1985, p. 313) définit l'œuvre comme un « principe vide qui aurait organisé la réalité extratextuelle pour que le lecteur puisse constituer un monde impossible à déterminer à partir des seuls éléments donnés ». Cela signifie que le texte ne fournit pas un sens déjà construit. Il propose plutôt des éléments, et c'est au lecteur de construire lui-même une certaine cohérence. Du coup, la lecture est peut-être moins stable qu'on l'imagine — et parfois aussi un peu plus fragile.

Ces « blancs », ces « discontinuités textuelles », ne sont pas des défauts à corriger. Au contraire, ils « stimulent l'activité de représentation du lecteur » (Iser, 1985, p. 322) et l'obligent à intervenir activement. Là où Jauss insistait sur l'ajustement

entre horizons, Iser introduit une exigence différente, parfois en tension avec la première : il faut du vide, de l'indétermination, sans quoi il n'y a plus vraiment lecture, mais simplement transmission d'informations.

La simplification pose cependant une difficulté. Rendre un texte plus clair est souvent indispensable avec des apprenants débutants. Pourtant, chaque fois qu'un implicite est expliqué ou qu'une ambiguïté disparaît, le lecteur dispose de moins de possibilités pour construire sa propre interprétation. L'adaptation facilite donc la compréhension, mais elle modifie aussi, au moins en partie, l'expérience de lecture.

Face à cette tension, on ne peut pas vraiment ignorer l'impasse pédagogique qui se dessine. Ajuster les horizons reste nécessaire pour rendre le texte accessible ; mais en même temps, préserver des zones d'indétermination l'est tout autant si l'on veut maintenir une véritable activité de lecture. On peut alors se demander, sans réponse totalement satisfaisante d'ailleurs, si cette tension est réellement tenable dans les pratiques de FLE. Simplifier revient donc à marcher sur une ligne assez instable. Trop de clarté fait perdre au texte sa dynamique interne ; trop d'opacité empêche la lecture d'avoir lieu. Il n'existe pas de solution unique, mais une série de choix, souvent dépendants du contexte.

La réflexion d'Eco (1965, p. 9) vient encore déplacer les lignes de cette discussion. Pour lui, un texte littéraire ne livre pas un sens unique. Plusieurs significations peuvent y être présentes en même temps, sans qu'il soit nécessaire d'en choisir une seule. Cette ouverture fait partie de la nature même de la littérature.

Le texte ne propose pas un sens unique que le lecteur n'aurait plus qu'à retrouver. Il laisse subsister plusieurs possibilités d'interprétation, sans en privilégier une seule. Cette idée prolonge la réflexion d'Iser tout en la déplaçant légèrement. Les « blancs » ne correspondent plus seulement à des éléments que le lecteur doit compléter ; ils participent aussi à l'organisation générale du texte et à l'ouverture de ses significations.

Dans une classe de FLE, cette ouverture peut néanmoins représenter une difficulté, notamment pour des apprenants débutants qui recherchent souvent une interprétation plus stable. Eco (1965, p. 23) évoque à ce propos une « réserve inépuisable de

significations », rappelant qu'aucune lecture ne peut épuiser complètement le sens d'une œuvre. Or, la simplification conduit fréquemment à réduire cette pluralité. En explicitant certains implicites ou en privilégiant une interprétation, elle limite parfois les différentes lectures que le texte original rendait possibles.

En définitive, les perspectives de Jauss, Iser et Eco convergent vers une idée centrale : le sens n'est pas une donnée immanente, mais une actualisation dont le lecteur constitue le pivot. Que l'on insiste sur la rencontre des horizons chez Jauss, les « blancs » du texte chez Iser ou l'ouverture presque infinie chez Eco, lire devient avant tout un processus où plusieurs sens peuvent coexister, sans qu'un seul ne s'impose définitivement.

Dans le cadre de cette recherche, cette idée invite à envisager la simplification autrement. Elle ne peut pas être réduite à une simple opération de clarification. Il s'agirait plutôt d'un équilibre assez fragile à construire. Simplifier un texte pour un niveau A2, ce n'est pas seulement enlever les difficultés linguistiques, mais aussi veiller à ne pas fermer complètement les possibilités de lecture. En ce sens, une simplification pertinente pourrait consister à alléger le texte sans l'enfermer, tout en laissant un espace minimal d'interprétation. Lorsque tout devient trop explicite, le texte risque en effet de perdre une dimension essentielle de lecture : celle de la co-construction du sens par le lecteur.

1.1.3. Médiation pédagogique et plaisir de lire

L'entrée en littérature, surtout lorsqu'elle s'adresse à un public dont la langue n'est pas l'ancrage maternel, ne saurait se réduire à une simple mise en présence, presque miraculeuse, de l'œuvre et de l'élève. Dans cette rencontre, il y a une distance, une zone d'ombre qu'il appartient à la pédagogie de combler, sans pour autant l'annuler totalement. Longtemps, on s'est accommodé d'une figure de l'enseignement qui se confondait avec celle de détenteur exclusif d'un savoir, chargé de verser une interprétation autorisée dans un récipient passif. Ce modèle est aujourd'hui frappé d'obsolescence, car il ignore la dynamique propre à la réception. Comme le note Laurent (2008, p. 25), le paysage didactique a vu « le modèle du maître transmetteur être remplacé par le maître-médiateur, le maître qui accompagne les apprentissages ».

Cette mutation n'est pas qu'un simple changement de terminologie ; elle redéfinit l'acte d'enseigner comme un geste de compagnonnage, où le sens n'est plus donné d'avance, mais cherché de concert.

Dans cette épaisseur de la médiation, le concept d'« étayage » forgé par Bruner (1983) s'impose comme la clé de voûte de notre réflexion. Pour Bruner, étayer, c'est mettre en place une structure temporaire, un échafaudage cognitif qui « rend l'enfant ou le novice capable de résoudre un problème, de mener à bien une tâche ou d'atteindre un but qui auraient été sans cette assistance, au-delà de ses possibilités » (Bruner, 1983, p. 263). Appliquée au texte littéraire, et plus particulièrement à la prose de Zola, cette idée conduit à envisager l'adaptation autrement. Il ne s'agit plus uniquement de simplifier le texte ou d'en réduire certains aspects stylistiques, mais de faciliter un accès progressif à sa compréhension. L'adaptation peut ainsi servir de médiation en aidant les apprenants à surmonter certaines difficultés linguistiques et interprétatives au cours de la lecture. En son absence, le texte littéraire risque au contraire de devenir, pour les apprenants débutants, un objet de lecture excessivement dense, où la charge linguistique et interprétative dépasse les capacités de traitement disponibles, compromettant l'accès au sens global et à l'expérience de lecture.

Ce rôle de médiateur n'est pourtant pas une simple fonction de simplification technique que l'on pourrait déléguer sans garde-fou. Il engage une responsabilité humaine et éthique que l'UNESCO (2024) vient opportunément rappeler dans un paysage saturé par des outils numériques qui pourraient laisser croire à une possible automatisation de la transmission culturelle. L'institution insiste sur l'urgence de « protéger les droits des enseignants [...] mais aussi la valeur de leurs pratiques », en rappelant que ce sont eux qui « facilitent la réflexion complexe, organisent les interactions des individus et promeuvent les valeurs humaines » (UNESCO, 2024, p. 26). Il y a là un point de rupture fondamental avec la génération probabiliste des machines : le médiateur humain, contrairement à l'algorithme, n'offre pas seulement une version plus courte ou plus simple du texte ; il lui insuffle une direction intentionnelle, il crée un espace où le texte « fait sens » pour un individu précis, dans une culture donnée.

C'est dans cette même lignée que Numa-Bocage (2007, p. 57) définit la médiation comme l'ensemble des démarches par lesquelles l'enseignant accompagne progressivement l'élève afin de l'aider à comprendre et à construire ses propres connaissances. La médiation n'est donc pas située dans le texte simplifié, elle réside dans le processus de sa manipulation. Elle est ce mouvement sinueux par lequel le sujet lecteur s'approprie ce qui, au départ, lui était radicalement étranger. Cette perspective trouve une résonance particulière dans les recherches du professeur Hoàng Văn Vân (2007) qui, analysant le contexte de rénovation éducative, démultiplie les facettes de cette fonction :

Thứ nhất là vai trò của người tổ chức nguồn lực và chính mình là nguồn lực. Thứ hai là vai trò của người hướng dẫn các quy trình và các hoạt động trong lớp học. [...] Vai trò thứ ba của người giáo viên trong đường hướng lấy người học làm trung tâm là người nghiên cứu và người học.

(Hoàng Văn Vân, 2007, p. 120)

Le premier rôle est celui d'organisateur des ressources et l'enseignant est lui-même une ressource. Le second est celui de guide des processus et des activités en classe. [...] Le troisième rôle de l'enseignant, dans l'approche centrée sur l'apprenant, est celui de chercheur et d'apprenant.

(Notre traduction)

Pour lui, l'enseignant est à la fois « celui qui organise les ressources », une « ressource en lui-même », mais aussi, et c'est peut-être le plus crucial, un « chercheur et un apprenant ». En se positionnant comme chercheur de sa propre pratique, l'enseignant qui adapte le naturalisme de Zola ne se contente pas d'appliquer une recette de lisibilité ; il éprouve les limites du compréhensible et les seuils de l'émotion. Il tâtonne.

Cependant, il faut bien admettre que la médiation pédagogique ne saurait être un simple exercice de transfert d'informations, aussi bien étayé soit-il. Comme le souligne avec insistance Lebrun (1993, p. 111), l'école a trop souvent privilégié des « cheminements stéréotypés » au détriment d'une « véritable ouverture sur l'univers

imaginaire ». Cette idée rejoint d'ailleurs une perspective bien établie en didactique, selon laquelle la lecture constitue un « espace de plaisir et de liberté qui invite à l'épanchement de l'affectivité, de la sensibilité, et au déploiement de l'imaginaire » (Cuq & Gruca, 2005, p. 427). Ce constat est inquiétant : à force de vouloir tout baliser, on risque de tuer l'objet même que l'on souhaite transmettre. Pour Lebrun, l'enjeu est vital, car « les stratégies de lecture à développer doivent toucher tout autant le cœur que la raison, sous peine de provoquer la désaffection du livre » (Lebrun, 1993, p. 115). Cela fait fortement écho au cri du cœur déjà ancien de Daniel Pennac (1992) : si le plaisir n'est pas le moteur, l'apprentissage devient une contrainte aride. Et on retrouve cette même opposition entre plaisir et contrainte chez Barthes (1973, p. 13), qui définit le plaisir comme « celui qui contente, emplit, donne de l'euphorie ; celui qui vient de la culture, ne rompt pas avec elle, est lié à une pratique confortable de la lecture ».

Il existe en effet un risque réel : celui d'une médiation si lourde, si explicative, qu'elle finirait par étouffer la liberté du lecteur. À force de simplifier, d'étayer, on pourrait oublier que la littérature est d'abord un espace de gratuité. Pennac (1992, p. 64) nous met en garde contre cette dérive utilitaire en érigeant « le droit de ne pas lire » en principe souverain. Pour Pennac, le plaisir ne se décrète pas, il ne se commande pas. On ne peut pas forcer un élève à aimer Zola par la seule vertu d'un texte rendu facile. On peut seulement, avec humilité, créer les conditions d'un désir possible.

C'est là que le lien avec notre problématique de la simplification devient central. Si le texte original se dresse comme un mur infranchissable, il n'y a tout simplement pas de plaisir possible, seulement la souffrance du déchiffrement aride. Une forme d'ajustement du texte apparaît alors comme une condition nécessaire pour rendre la lecture praticable et permettre l'entrée dans l'univers littéraire. Elle vise, *in fine*, à maintenir la motivation du lecteur. Simplifier, ce n'est pas trahir le projet naturaliste, c'est au contraire faire le pari que l'univers de Zola est assez puissant pour survivre à une réduction de sa complexité de surface. En rendant le texte abordable, le médiateur permet à l'apprenant d'exercer son métier de lecteur sans épuiser ses ressources vitales dans une lutte ingagnable contre le lexique textile du XIX^e siècle.

Le plaisir de lire, loin d'être un simple élément décoratif ou un bonus pédagogique, est la condition importante d'une rencontre réussie avec l'altérité textuelle. Sans lui, l'étayage décrit par Bruner perd sa dimension dynamique et se réduit à un dispositif purement technique, dépourvu d'adhésion du lecteur. Avec lui, la simplification peut être envisagée comme un dispositif de médiation permettant au sujet lecteur de niveau A2 d'accéder progressivement au texte de Zola, non pas seulement comme objet d'étude, mais comme expérience de lecture impliquant interprétation, projection et engagement subjectif. Le médiateur — qu'il s'appuie sur son expertise humaine ou qu'il utilise de manière critique les potentialités d'une IA — reste le garant de cet équilibre précaire entre l'exigence d'accessibilité et la persistance de la beauté. C'est dans ce dialogue tendu entre les structures d'aide et les espaces de pure gratuité que se joue l'avenir de la littérature en classe de langue.

1.2. Le défi de l'accessibilité : théories de la simplification des textes

1.2.1. L'input compréhensible et la médiation linguistique

L'exposition à un texte, si authentique soit-il, ne suffit pas, à elle seule, à garantir une acquisition linguistique. On pourrait être tenté de croire qu'il suffit de lire pour apprendre, comme si le sens se déposait naturellement dans l'esprit du lecteur. Pourtant, dans le cas d'un apprenant allophone, la lecture ressemble moins à une réception immédiate qu'à un processus de compréhension qui mobilise des ressources cognitives limitées. Autrement dit, ce n'est pas seulement le texte qui compte, mais la manière dont il peut être traité. Dans cette optique, la difficulté de l'« input » ne peut pas être réduite à une question de motivation ou de goût personnel ; elle agit plutôt comme une contrainte qui influence fortement la possibilité même d'engager une activité de compréhension.

L'hypothèse fondamentale de Krashen (2013) consiste à poser que l'acquisition n'est possible que si l'apprenant est exposé à un message légèrement au-dessus de son niveau actuel. Il formule ce principe à travers le concept du « $i+1$ » :

For an item of grammar to be acquired, the language acquirer must be ready to acquire the item. It must, in other words, be at the acquirers' $i+1$, where i = aspects of grammar that were most recently acquired.

(Krashen, 2013, p. 102)

Pour qu'un point de grammaire soit acquis, l'acquéreur de la langue doit être prêt à l'acquérir. Il doit, en d'autres termes, se situer au niveau $i+1$, où i représente les aspects de la grammaire acquis le plus récemment.

(Notre traduction)

Si l'on essaie de transposer cette idée à la lecture littéraire, une question se pose presque immédiatement. Dans le cas d'un apprenant de niveau A2, la rencontre avec une œuvre originale ne correspond que rarement à ce léger décalage décrit par Krashen. On pourrait même dire, de manière volontairement imagée, qu'il s'agit parfois d'un " $i+10$ ". Ce décalage ne se limite pas à quelques mots inconnus : il concerne à la fois la densité lexicale, les implicites culturels et la complexité syntaxique. À partir d'un certain seuil, le texte ne devient pas simplement difficile, il tend à échapper à toute prise. Ce serait peut-être à ce moment précis que quelque chose se bloque : l'apprenant ne lit plus vraiment, il tente de décoder. Le texte, au lieu de soutenir une construction de sens, risque alors de devenir une suite d'éléments fragmentés, difficilement articulables entre eux.

Cette idée se retrouve, sous une autre forme, dans les travaux de Vygotsky (1978) sur la « Zone Proximale de Développement » (ZPD). Là où Krashen insiste sur la qualité de l'input, Vygotsky met l'accent sur ce que l'apprenant peut accomplir avec un accompagnement adapté. Il définit cette zone comme un espace de maturation :

The zone of proximal development defines those functions that have not yet matured but are in the process of maturation, functions that will mature tomorrow but are currently in an embryonic state. These functions could be termed the “buds” or “flowers” of development rather than the “fruits of development.

(Vygotsky, 1978, p. 86)

La zone de proche développement définit les fonctions qui ne sont pas encore arrivées à maturité mais qui sont en cours de maturation, des fonctions qui arriveront à maturité demain mais qui sont actuellement dans un état embryonnaire. Ces fonctions pourraient être qualifiées de “bourgeons” ou de “fleurs” du développement plutôt que de “fruits” du développement.

(Notre traduction)

Ce rapprochement entre Krashen et Vygotsky semble particulièrement fécond, même s'il n'est pas entièrement évident à stabiliser. D'un côté, si l'input est trop éloigné, il ne peut pas être intégré ; de l'autre, un certain dépassement reste nécessaire pour que le développement ait lieu. La difficulté consiste donc à situer ce point d'équilibre, qui n'est jamais donné à l'avance. Lorsqu'un texte dépasse trop largement les capacités de traitement de l'apprenant, les « bourgeons » évoqués par Vygotsky ne disposent pas des conditions nécessaires pour se développer. L'activité de lecture risque alors de se réduire à un décodage mot à mot, coûteux et peu efficace. On peut même supposer que la mémoire de travail se trouve rapidement saturée dans ce type de situation : toute l'attention est mobilisée par le déchiffrage, au détriment de la construction d'une représentation globale du texte. Et c'est peut-être là que se joue une forme de rupture, non pas au sens absolu, mais comme un glissement progressif vers une lecture qui n'en est plus vraiment une. Ce constat, en réalité, paraît assez

proche de ce que l'on observe empiriquement chez de nombreux apprenants, même si on a parfois tendance à le sous-estimer.

Dans cette perspective cognitive, Kintsch (1982, p. 777) considère que la compréhension d'un texte ne peut pas être réduite à l'application de règles prédéfinies. Les lecteurs ajustent leurs stratégies au fil de la lecture pour construire le sens. Celles-ci les orientent dans leur interprétation, sans garantir pour autant un résultat identique d'un lecteur à l'autre. Encore faut-il, cependant, que le texte offre un minimum d'ancrage pour que ces stratégies puissent se déployer. Sans cela, l'inférence devient presque impossible, ou du moins extrêmement fragile.

C'est à partir de ce point que la question de la médiation apparaît comme difficilement évitable. Dans le cadre de cette recherche, il ne s'agit pas de penser la médiation comme une interaction entre individus, mais comme une intervention sur le texte lui-même. L'objectif n'est pas de simplifier le texte pour lui-même, mais de le rendre plus accessible à un apprenant donné. Cela peut impliquer un vocabulaire plus fréquent, des référents plus explicites ou des structures syntaxiques moins complexes. Ces ajustements cherchent à maintenir un niveau de difficulté compatible avec les connaissances de l'apprenant, dans une logique proche du « $i+1$ » de Krashen et de la zone proximale de développement de Vygotski. Cet ancrage théorique servira de repère pour analyser et comprendre les différentes formes de simplification dans la suite de notre étude.

Cependant, cette transformation du texte n'est jamais totalement neutre. Blum-Kulka (1986, p. 21) avance que « l'explicitation est une stratégie universelle inhérente au processus de médiation linguistique ». Ce point semble important, car il montre que rendre un texte plus explicite n'est pas une opération artificielle, mais un mécanisme profondément ancré dans la communication. Cette situation met en évidence une difficulté. Plus le texte est explicité, plus les implicites se réduisent, ce qui limite aussi une partie des possibilités d'interprétation. L'adaptation facilite donc l'accès au texte, mais elle modifie en même temps l'expérience de lecture.

La simplification textuelle peut ainsi être considérée comme un moyen de réduire l'écart entre les exigences du texte et les ressources des apprenants, sans pour autant

le faire disparaître complètement. Elle facilite l'entrée dans la lecture, mais elle ne suffit pas, à elle seule, à garantir un véritable travail d'interprétation. Autrement dit, rendre un texte accessible constitue sans doute une condition nécessaire, mais certainement pas suffisante. Et c'est peut-être dans cet espace, encore un peu incertain, que se situe l'enjeu réel de la didactique de la lecture.

1.2.2. Définition et niveaux de simplification

Avant même d'être formalisée comme objet théorique, la simplification textuelle s'inscrit dans une pratique éditoriale déjà ancienne. Collès (2002, p. 402) situe l'apparition des premiers textes en français simplifié à la fin des années 1960. Leur développement est étroitement lié à la collaboration entre la maison Hatier et Philippe de Beaumont. Dès leur création, ces ouvrages répondent à un objectif concret : proposer des textes plus accessibles aux apprenants. La simplification s'inscrit ainsi dans une démarche éditoriale destinée à faciliter la lecture, en particulier dans le contexte de l'enseignement du FLE.

C'est en travaillant sur les niveaux d'analyse que la question de la simplification devient réellement moins stable qu'elle n'en a l'air. L'auteur propose ainsi que les critères permettant « d'arriver à une version plus brève et plus accessible » peuvent être « regroupés selon quatre niveaux : lexical, syntaxique, stylistique et culture » (Collès, 2002, pp. 404-410). Cette liste semble claire au premier regard. Pourtant, plus on avance dans l'analyse des textes littéraires, plus cette séparation paraît artificielle, comme si les catégories ne tenaient pas complètement ensemble.

Cette remise en question ouvre la question du statut même du niveau culturel qui, tel qu'il est envisagé dans la typologie de Collès (2002), fait l'objet de débats dans la littérature en raison de la difficulté de sa délimitation et de son caractère non stabilisable comme catégorie d'analyse autonome (Charaudeau, 2009). Il est généralement considéré comme une dimension contextuelle du discours (Maingueneau, 2004) plutôt que comme un niveau d'analyse strictement autonome. Cette difficulté souligne le caractère diffus et transversal de la culture, qui traverse le texte sans se laisser facilement circonscrire dans une catégorie stable.

Face à cette hétérogénéité des niveaux d'analyse, les approches plus opérationnelles tendent à recentrer l'analyse sur des transformations observables du texte. Buet (2022) propose une structuration plus opérationnelle :

[...] différentes opérations peuvent être considérées pour arriver à ce résultat : remplacement des mots compliqués ou techniques (simplification lexicale), restructuration grammaticale des phrases (simplification syntaxique), schématisation des idées (simplification conceptuelle), redondance et explicitation des points clefs (modification élaborée), suppression des informations secondaires pour réduire la longueur et faire ressortir l'essentiel (résumé ou compression).

(Buet, 2022, p. 15)

Ce qui paraît important ici, ce n'est pas seulement la liste, mais le fait que la simplification cesse d'être une idée abstraite. Elle devient une série de gestes, presque mécaniques parfois, mais pas totalement cohérents entre eux. La présence de la « modification élaborée », avec « redondance » (Buet, 2022) et « explicitation » (Blum-Kulka, 1986), casse d'ailleurs une intuition assez classique : simplifier ne veut pas toujours dire enlever. Parfois, on ajoute, et c'est là que le modèle se complique. Avant d'arriver à cette vision opérationnelle, il faut pourtant passer par une autre couche, plus théorique mais aussi plus ancienne. Historiquement, Chall et Dale (1995, cités par Brouwers et al., 2012, p. 212), montrent que les obstacles à la compréhension viennent surtout du « lexique et de la structure syntaxique ». C'est une vision assez directe, presque rassurante : si on simplifie les mots et les phrases, le texte devient accessible. Mais cette idée s'effondre un peu dès qu'on regarde des textes littéraires, où la compréhension ne dépend pas uniquement de la langue visible. C'est probablement pour cette raison que Saggion (2017, cité par Gala et al., 2018, p. 123), propose autre chose : la simplification comme « l'adaptation d'un texte afin de le rendre plus accessible à un public ciblé, tout en conservant le contenu sémantique intact ». Ici, le déplacement est important. On ne touche plus seulement

à la structure linguistique, mais à une logique de médiation. Le sens devient une sorte de frontière implicite : on peut transformer la forme, mais pas franchir ce seuil.

En réalité, cette tension apparaît déjà si on revient en arrière, vers Causa (2001, p. 65-69), qui affirme que « la simplification se situe au niveau du système linguistique » et qu'elle consiste essentiellement en une « réduction du système linguistique ». Cette position paraît presque radicale aujourd'hui, surtout au regard des approches plus récentes (comme celles de Saggion, 2017 ou de Buet, 2022), qui envisagent la simplification comme une opération d'adaptation ou de transformation du discours, mais elle conserve une cohérence interne : elle suppose que simplifier revient à diminuer la complexité structurelle elle-même. Ce point de vue reste dérangeant, surtout si on le confronte à des objets esthétiques où la complexité fait partie du sens. Ce contraste entre réduction et adaptation ne se limite pas à une simple opposition théorique : il reflète des choix méthodologiques différents selon les approches. Il crée en fait un espace instable dans lequel toutes les autres approches vont se situer. La simplification devient alors quelque chose qui oscille, sans jamais vraiment se fixer. Au final, la simplification ne peut pas être définie de manière totalement stable. Elle peut ainsi être envisagée comme un ensemble d'opérations portant sur différents niveaux du texte, sans que ces niveaux soient toujours strictement séparables.

À partir de là, deux niveaux d'analyse peuvent être distingués de manière fonctionnelle plutôt que hiérarchique. Le premier est linguistique, couvrant le lexique et la syntaxe, correspondant à la surface visible du texte. Le second est discursif, incluant l'organisation, la hiérarchisation de l'information et la circulation du sens dans l'ensemble du texte. Le niveau culturel est ici considéré comme une dimension contextuelle transversale, difficilement isolable en tant que catégorie d'analyse autonome.

1.2.3. Les opérations de simplification

Après avoir défini les différents niveaux de simplification, il nous faut s'intéresser aux opérations qui permettent de transformer concrètement les textes. La simplification textuelle ne peut pas être réduite uniquement à une simple diminution de la difficulté. On pourrait avoir l'impression qu'il suffit d'effacer ce qui gêne, mais

le processus est en réalité plus complexe. La littérature montre bien que ce processus est loin d'être homogène. Il s'appuie sur une pluralité d'opérations que les chercheurs tentent de classer, non seulement pour mieux les décrire, mais aussi pour comprendre leurs effets réels sur la compréhension.

Dans cette perspective, Hijazi (2020) insiste sur le fait que la simplification ne constitue pas une action unique, mais un ensemble structuré de transformation. Elle identifie ainsi les classes principales d'opérations, ce qui oblige à dépasser une vision trop simpliste du phénomène. Parmi ces opérations, une se détache nettement : le découpage des phrases, décrit comme « l'opération la plus fréquente » (Hijazi, 2020, p. 144). Ce point mérite qu'on s'y attarde. Découper une phrase longue en unités plus courtes ne consiste pas seulement à alléger la syntaxe. Cela modifie aussi la manière dont l'information est distribuée. Pour un apprenant de FLE, cela peut faire toute la différence, car la compréhension dépend souvent de la capacité à identifier rapidement les relations entre les éléments. Or, dans une phrase complexe, cette capacité est facilement mise à l'épreuve. Le découpage agit donc comme un outil de guidage, presque comme une aide à la lecture intégrée au texte lui-même.

D'autres travaux adoptent une approche plus resserrée. L'étude de Gala et al. (2020, p. 9), par exemple, choisit de se concentrer sur la dimension syntaxique et propose une « typologie spécifique » fondée sur trois opérations principales : « des substitutions, des suppressions et des découpages ». Cette triade a l'avantage d'être directement mobilisable dans l'analyse. La suppression correspond à l'effacement d'éléments jugés secondaires. Elle peut concerner des segments très courts, comme des adjectifs, mais aussi des portions plus larges du texte. Ce qui semble intéressant ici, c'est que la suppression ne se limite pas à un simple raccourcissement du texte. Elle modifie aussi ce qui est perçu comme essentiel. La substitution, de son côté, repose sur un remplacement. On ne supprime pas, on reformule à l'intérieur d'un cadre plus restreint, souvent en choisissant un mot plus fréquent ou une structure plus simple. Quant au découpage, déjà mentionné, il transforme la structure globale de la phrase. Ces trois opérations ne sont pas équivalentes. Elles agissent à des niveaux différents et produisent des effets distincts, même si, dans la pratique, elles sont

souvent combinées. Il convient également de préciser que ces opérations ne se rattachent pas à un seul niveau de manière exclusive. Une même opération peut intervenir au niveau lexical, syntaxique ou discursif, selon le type de transformation effectué. Il reste toutefois une tension difficile à éviter : simplifier implique presque toujours une forme de perte, qu'elle soit sémantique ou stylistique. Et cette perte n'est jamais totalement neutre. Elle apparaît de manière particulièrement sensible dans l'écriture de Zola, où l'accumulation d'adjectifs ne sert pas uniquement à décrire, mais contribue à créer une forme de pression du milieu sur les personnages. Leur suppression, si elle allège la phrase sur le plan syntaxique, tend aussi à modifier en partie cette atmosphère dense et parfois étouffante, caractéristique du naturalisme.

La simplification ne se limite pourtant pas à ces ajustements visibles. Elle peut aller plus loin, vers une transformation plus profonde de l'énoncé. C'est ce que montrent les travaux de Fuchs (2020a), qui abordent la question sous l'angle de la paraphrase et de la reformulation. Simplifier, dans ce cadre, consiste à produire un énoncé équivalent, mais plus accessible. Cela rejoint la réflexion de Martinot (2010), pour qui la reformulation joue un rôle central dans la médiation linguistique. Dire autrement ne signifie pas seulement simplifier la forme ; cela implique parfois de modifier le point de vue, d'explicité certaines relations, ou de réorganiser l'information. À ce stade, on n'est plus dans une transformation locale, mais dans une reconfiguration plus globale du discours.

Cette logique conduit presque naturellement à la notion d'explicitation, théorisée par Blum-Kulka (1986, p. 21), selon laquelle, « explicitation is a universal strategy inherent in the process of language mediation » (l'explicitation est une stratégie universelle inhérente au processus de médiation linguistique – Notre traduction). Comme nous l'avons déjà suggéré précédemment, l'idée peut sembler paradoxale au départ. Simplifier ne signifie pas toujours réduire. Il arrive au contraire qu'il faille ajouter des éléments : ajouter un lien logique, préciser un référent, lever une ambiguïté. Dans le contexte du FLE, ce type d'intervention paraît particulièrement important. Certains implicites, évidents pour un lecteur natif, deviennent de véritables obstacles pour un apprenant. L'explicitation rend alors le texte plus lisible, même si

cela passe par une forme d'allongement. Ces différentes approches permettent ainsi de dégager un cadre théorique commun, dans lequel la simplification est envisagée comme un ensemble d'opérations complémentaires, telles que la suppression, la substitution, la reformulation et l'explicitation.

Pour conclure, on pourrait dire que ces différentes opérations ne fonctionnent pas de manière isolée. Elles s'entrecroisent, se superposent, parfois même se confondent. Une suppression peut s'accompagner d'une reformulation. Une substitution peut entraîner une explicitation. Il devient alors difficile de tracer des frontières nettes. Cette instabilité n'est pas forcément un problème. Elle reflète plutôt la complexité du travail de simplification lui-même. Trouver un équilibre entre réduire et expliquer, alléger sans appauvrir, guider sans dénaturer — tout cela reste une opération assez fragile. Et c'est précisément dans cet espace, un peu incertain, que se joue la médiation des textes littéraires en FLE.

1.3. Simplifier Zola : enjeux de littérarité et pertes stylistiques

1.3.1. Le projet naturaliste de Zola

Comprendre l'écriture de Zola demande de ne pas s'arrêter à l'étiquette du naturalisme. Ce qui se joue derrière est moins une question de style qu'une certaine manière de penser le roman. Dans *Le Roman expérimental*, Zola (1880) ne cherche pas d'abord à être original, mais à rapprocher la littérature d'une forme d'observation du réel. Le détail, chez lui, n'est pas décoratif. Il sert à produire un effet de vérité, presque comme une preuve.

Cette idée d'écriture proche de la science apparaît encore plus clairement si l'on considère que, pour Zola, la littérature devient une véritable « expérience ». Il ne s'agit plus d'imaginer librement, mais de suivre une forme de méthode. La remarque de Matthews (1957, p. 20) éclaire bien ce déplacement : « la Science est “de la poésie expliquée”, et le savant, “un poète qui remplace les hypothèses de l'imagination par l'étude exacte des choses et des êtres.” » On voit alors se dessiner une tension assez intéressante : écrire, oui, mais en limitant volontairement ce que l'imaginaire pourrait ouvrir. Dit autrement, une première forme de réduction épistémologique est déjà à

l'œuvre, c'est-à-dire une manière de limiter volontairement les variables du réel, même si elle ne se formule pas encore comme une simplification au sens pédagogique. Cette conception entraîne alors un déplacement important : l'écrivain naturaliste se rapproche d'un observateur. Zola s'inscrit dans un contexte marqué par le positivisme, mais il ne se contente pas de copier le réel. Il organise, il choisit, il met en évidence certains éléments plutôt que d'autres. Comme le souligne Mitterand (1985, p. 187), « le discours naturaliste de Zola renoue bien avec l'héritage de l'esprit des Lumières [...] pour revendiquer la liberté du jugement et la liberté de la création. » Il y a donc une forme de tension entre rigueur et liberté, qui rend l'ensemble moins mécanique qu'on pourrait le croire au premier regard.

Cette logique devient encore plus visible lorsqu'on observe la place accordée aux déterminismes. Dans la préface de la deuxième édition de *Thérèse Raquin*, Zola (1868/2012) affirme : « j'ai voulu étudier des tempéraments et non des caractères [...], dominés par leurs nerfs et leur sang ». Ce choix n'est pas anodin. Il revient à réduire la complexité du sujet humain pour le rendre plus lisible, presque analysable. On n'est pas encore dans une simplification au sens didactique, mais il y a déjà ce geste qui consiste à rendre le réel plus saisissable.

Représenter le monde tel qu'il est implique alors une certaine économie : tout ne peut pas être gardé, tout ne doit pas être dit de la même manière. C'est peut-être là que le lien avec la simplification devient plus concret. Avant même toute adaptation pédagogique, le texte naturaliste opère déjà des choix, des réductions, des focalisations. Et c'est précisément cette organisation du texte, pensée pour rendre le réel plus lisible, qui peut produire une forme d'apparente opacité pour certains lecteurs, notamment en contexte d'apprentissage, ce qui explique en partie les difficultés rencontrées par les apprenants.

1.3.2. Le style zolien comme système

À la première lecture, il est difficile de saisir clairement ce qui fait la particularité de l'écriture de Zola. On a pourtant l'impression que le texte est organisé par quelque chose de plus profond, comme un mécanisme discret qui structure l'ensemble. L'écriture de Zola ne fonctionne pas vraiment comme un simple décor qu'on ajoute

sur une intrigue. L'idée que les descriptions ralentissent le récit ne tient pas vraiment, elles structurent plutôt le texte. On retrouve cette logique chez Hamon (1993, p. 5), lorsqu'il évoque « l'hypertrophie de certaines opérations générales comme [...] la redondance, [...] la mise en facteur commun ». Autrement dit, la description n'est pas un ajout, mais une manière de construire le texte. Le choix des mots et leur place dans la phrase participent ensemble à cette construction. Dans le même sens, Clamens (1959, p. 522) rappelle que la description ne dépend pas uniquement des mots employés. Sa structure participe aussi au choix du lexique et à la place que chaque mot occupe dans la phrase. Chez Zola, la description ne sert donc pas seulement à montrer. Elle entre aussi en jeu dans l'organisation même du texte.

Cette volonté de tout dire se remarque assez vite. Le texte accumule, mais cette accumulation n'est pas gratuite ni simplement décorative. Selon Mansouri (2025, p. 147), l'accumulation ne relève pas seulement d'un choix stylistique. Elle fait partie de la manière dont Zola organise ses phrases pour donner forme au réel. Cette lecture aide à comprendre pourquoi les phrases de Zola avancent par ajouts successifs. Les détails s'ajoutent un par un, ce qui donne au texte une forte densité. Par moments, cela donne une impression assez claire et visuelle, mais cette accumulation peut aussi alourdir la lecture, et l'effet produit n'est pas toujours totalement maîtrisé du point de vue du lecteur.

La question du rythme peut également entraîner une certaine difficulté de compréhension chez le lecteur. Les phrases sont assez longues, mais pas forcément fluides. Dans cette perspective, Brunet (2016, p. 3I.4) note que « la segmentation est donc plus serrée chez Zola que chez les autres écrivains. » En lisant Zola, on sent bien cette fragmentation. La ponctuation coupe souvent, au point qu'il faut reprendre, parfois relire. Brunet (2016, p. 3I.12) parle aussi des « grandes fresques au rythme large, du *Ventre de Paris* à la *Débâcle* en passant par *Germinal* et la *Terre* ». Ces passages donnent une impression d'ampleur, mais la lecture reste exigeante. La longueur ne signifie pas lenteur, puisqu'il y a en réalité beaucoup d'informations, très vite.

Le lexique renforce cette impression de densité et de fragmentation du texte. Zola ne reste pas dans la vague ; il nomme. Comme l'explique Brunet (1985) :

[...] son lexique s'en trouve diminué d'autant, ce qui est paradoxal dans une œuvre où foisonnent les objets, où l'on trouve toutes les variétés de fleurs, d'arbres, de fromages, de poissons, de viandes, d'outils, d'étoffes, toute la technologie du train et de la mine, toutes les techniques commerciales, financières, agraires, tous les objets du culte et de l'art. En réalité son catalogue fait penser à une encyclopédie non à un dictionnaire de langue.

(Brunet, 1985, pp. 38, 39)

Pour les apprenants de FLE, cette caractéristique soulève un paradoxe. Ils rencontrent principalement un vocabulaire concret, mais en très grand nombre. Ce n'est pas forcément plus simple. Brunet (1985, p. 38) ajoute : « Mais Zola fuit l'abstraction [...] tandis que les suffixes qui désignent un agent ou un instrument concret sont en excédent ». Cette préférence pour le concret ne rend pas le texte plus accessible automatiquement. Elle change simplement la difficulté.

L'effet produit est assez fort. Cet effet peut être rapproché de l'analyse de Mansouri (2025, p. 146), qui parle de la syntaxe « convertissant l'accumulation grammaticale en immersion perceptive ». Cette idée correspond à ce que l'on peut ressentir dans certains passages : le texte donne une impression de présence. Cependant, cette immersion peut devenir lourde, car elle demande un effort constant.

Cet effet de présence et d'immersion ne se limite pas à une expérience de lecture : il soulève également des enjeux importants dans une perspective d'apprentissage. À ce propos, Dubois (1968, p. 258) souligne que le lecteur ne peut accéder à une « pleine jouissance esthétique » de l'œuvre littéraire. Cette expérience de lecture paraît toutefois difficile à atteindre pour des apprenants débutants en FLE. Dans le cadre de cette recherche, menée auprès d'apprenants de niveau A2, ce phénomène pourrait expliquer certaines réactions face aux extraits. Les difficultés ne se situent pas seulement au niveau du vocabulaire. La manière dont le texte est construit peut également compliquer la compréhension. La simplification cherche alors à rendre

cette organisation plus lisible, afin que les apprenants puissent suivre plus facilement le fonctionnement du texte. À défaut, celui-ci peut rester difficile à comprendre et à s'approprier.

1.3.3. Que perd-on en simplifiant Zola ?

Zola occupe une place à part, presque excessive, et cette impression de débordement se ressent très vite dès qu'on entre dans ses romans. Dans *Au Bonheur des Dames*, par exemple, la description du grand magasin ne fonctionne pas comme un simple décor. Dès les premières pages, avec l'arrivée de Denise et la découverte des rayons, des tissus, des vitrines qui débordent de soie et de couleurs, on a l'impression que le récit est absorbé par cet espace. Cette impression se vérifie très nettement dans la description même du magasin, où l'accumulation des marchandises envahit littéralement la page :

Cela partait de haut, des pièces de lainage et de draperie, mérinos, cheviottes, molletons, tombaient de l'entresol, flottantes comme des drapeaux, et dont les tons neutres, gris ardoise, bleu marine, vert olive, étaient coupés par les pancartes blanches des étiquettes. À côté, encadrant le seuil, pendaient également des lanières de fourrure, des bandes étroites pour garnitures de robe, la cendre fine des dos de petit-gris, la neige pure des ventres de cygne, les poils de lapin de la fausse hermine et de la fausse martre. Puis, en bas, dans des casiers, sur des tables, au milieu d'un empilement de coupons, débordaient des articles de bonneterie vendus pour rien, gants et fichus de laine tricotés, capelines, gilets... Denise vit une tartanelle à quarante-cinq centimes, des bandes de vison d'Amérique à un franc, et des mitaines à cinq sous. C'était un déballage géant de foire [...]

(Zola, 1883/2024, p. 35)

Le passage met en scène une logique d'accumulation presque ininterrompue d'objets — tissus, fourrures, articles de bonneterie — qui dépasse la simple fonction descriptive. Le texte ne se contente plus de représenter un lieu, il produit un effet de saturation qui entraîne le lecteur dans un espace presque envahissant. On ne suit plus

seulement une histoire, on circule à l'intérieur d'un univers saturé, quasiment oppressant. Chez Zola, la description ne se contente pas d'accompagner le récit : elle le prend en charge, jusqu'à devenir son moteur interne. Cette idée entre en résonance avec Étienne Brunet (2016), qui souligne :

[...] le goût de Zola pour la description le pousse à surenchérir sur son époque et à développer certains traits que le naturalisme avait hérités du romantisme. Personne n'ira plus loin dans ce sens et après Zola la littérature française rebrousse chemin.

(Brunet, 2016, p. 31.22)

Brunet suggère ici l'existence d'une tension déjà sensible chez Zola : celui-ci s'inscrit dans l'héritage du romantisme, tout en le dépassant, jusqu'à sembler repousser les limites du lisible.

Cette notion de limite devient justement problématique dès qu'on tente de simplifier ces textes pour les besoins de l'enseignement. Comme sur ce terrain, la difficulté ne tient pas seulement à la langue, mais à une densité qui résiste à la réduction. Dans certains passages de Zola, surtout dans *Germinal*, on a vraiment l'impression que si on enlève une seule description, quelque chose se casse. Par exemple, les longues scènes dans les mines ne sont pas « en plus » du récit : elles font partie du rythme même du texte. Du coup, quand on simplifie en classe, ce n'est pas juste une perte d'information. On sent parfois que le texte tombe à plat, comme si on avait retiré une pièce essentielle. C'est dans ce sens que le geste de simplification apparaît comme une opération risquée, même lorsqu'il est motivé pédagogiquement.

Cette tension est aussi soulignée par Gerolimich et Vecchiato (2014, p. 35), qui rappelle que « la forme dans laquelle un texte est rédigé peut aider ou contrarier la transmission d'un savoir au lecteur ». Leur remarque conduit à s'interroger sur le rôle de la forme dans la compréhension. La forme ne constitue pas seulement un support du contenu ; elle participe aussi à la construction du sens. Cette idée peut toutefois entrer en tension avec les pratiques de simplification en FLE, qui cherchent souvent à rendre les textes plus fluides. Une lecture plus fluide facilite généralement la

compréhension, mais elle laisse parfois moins de place aux moments où le lecteur est amené à interpréter et à élaborer le sens du texte.

Laidoudi et Messaour (2024) nuancent d'ailleurs cette intuition pédagogique :

[...] des séquences descriptives [...] fournissent de précieuses précisions qui permettent de comprendre pleinement ce qui est lu. Supposés facultatifs [...], les passages descriptifs ne pourraient être supprimés sans affecter la compréhension.

(Laidoudi & Messaour, 2024, p. 363)

Cette position entre en dialogue direct avec la pratique de simplification : ce qui semble secondaire dans une logique d'allègement est précisément ce qui structure la compréhension en profondeur. Autrement dit, l'accessoire n'est pas toujours accessoire — il peut devenir le lieu même de la compréhension différée.

Schreiber (2009, p. 96) radicalise encore cette idée en affirmant que « de petites coupures peuvent fausser le message du texte ». Mais ce qui est intéressant ici, c'est la continuité de sa réflexion avec celle d'autres chercheurs : là où Laidoudi et Messaour insistent sur la perte de compréhension, Schreiber insiste sur la transformation du message lui-même. Il ne s'agit plus seulement de comprendre moins, mais de comprendre autrement — donc potentiellement de comprendre de travers. Dans le cadre de la simplification, cette logique d'adaptation peut être rapprochée de certaines analyses de la traduction comme forme de transformation contrôlée du texte. Schreiber (2009, p. 98) ajoute ainsi qu'« il faudrait, à mon avis, indiquer qu'il ne s'agit pas d'une traduction dans le sens strict, mais d'une adaptation ». Cette précision introduit une exigence éthique : rendre visible la transformation opérée sur le texte.

Cette question de la transformation est encore accentuée par Laidoudi et Messaour (2024, p. 356) lorsqu'ils évoquent le risque de « déformer et dénaturer le texte original ». Ici, leur propos rejoint indirectement celui de Schreiber : ils pointent tous une zone fragile entre fidélité et adaptation, mais selon des angles différents — les analyses de Laidoudi et Messaour (2024) mettent en évidence les effets de la

suppression sur la compréhension et la fidélité au texte ; celles de Schreiber (2009) déplacent la réflexion vers la transformation du message lui-même.

Dans cette chaîne de réflexions, Reisinger (1987) introduit une autre perspective, presque opposée dans sa radicalité pédagogique :

Devant les textes aménagés de cette manière, le lecteur — enfant scolaire ou étudiant — finit par ne plus avoir besoin pratiquement de lire méticuleusement, car il n'y a presque plus rien à lire, à découvrir individuellement dans le texte.

(Reisinger, 1987, cité par Lebrun, 1993, p. 114)

Cette remarque entre en tension directe avec les positions précédentes : si les autres chercheurs alertent sur la déformation du texte, Reisinger met en cause la transformation du lecteur lui-même. Le problème ne serait donc plus seulement textuel, mais cognitif : un texte trop simplifié ne demande plus d'effort interprétatif, et c'est précisément cet effort qui disparaît.

Au fond, une question revient à travers les différentes positions évoquées : que devient un texte lorsqu'on cherche à le rendre accessible ? La simplification apparaît comme une évidence dans le cadre du FIE, mais elle n'est jamais totalement neutre. Quelque chose se modifie, quelque chose se perd aussi dans le passage d'un texte original à sa version simplifiée, même si cette perte reste difficile à cerner clairement. La question centrale de cette recherche est alors de savoir jusqu'où il est possible de simplifier les textes de Zola pour des apprenants de niveau A2, sans dénaturer l'expérience de lecture littéraire.

1.4. L'IA comme agent de simplification

1.4.1. Simplification éditoriale et simplification automatique

Si la médiation textuelle a longtemps été pensée comme un processus relativement simple — lire, comprendre, puis réécrire pour rendre un contenu accessible — cette représentation devient difficile à maintenir aujourd'hui, notamment lorsqu'on distingue la simplification éditoriale de celle générée par l'IA.

Dans la littérature, la simplification textuelle ne se limite pas à une seule pratique. Comme le souligne Delbart (1994, p. 133), on peut distinguer quatre méthodes

principales : la production de textes directement simplifiés pour un public spécifique, la réduction du texte d'origine par suppression d'éléments, la transformation du genre textuel (du roman à la scène de théâtre par exemple), et enfin la réécriture du texte initial sous une forme plus accessible. Parmi ces différentes formes, la présente recherche s'intéresse exclusivement à la réécriture de textes littéraires, qu'elle soit réalisée par une maison d'édition ou générée par ChatGPT.

Dans le cadre de ce travail, l'attention se porte principalement sur cette quatrième méthode, celle de la réécriture. Elle implique non seulement une transformation formelle du texte, mais aussi une adaptation au public visé, à ses compétences et à ses besoins, comme le suggèrent également les travaux en simplification textuelle de Bhatia (1983) ainsi que de Cardon (2021). Autrement dit, simplifier ne consiste pas seulement à raccourcir, mais à reconstruire un texte dans une logique de transmission. Dans la littérature relative au traitement automatique du langage, la simplification textuelle est souvent rapprochée d'une forme de traduction monolingue. Elle est alors modélisée comme une tâche de réécriture, dans laquelle les systèmes apprennent à produire des versions simplifiées à partir de corpus parallèles associant textes complexes et texte simplifiés (Daelemans et al., 2004 ; Zhu et al., 2010 ; Nisioi et al., 2017).

C'est précisément à ce niveau que la distinction entre simplification éditoriale et simplification automatique devient pertinente. Dans un cadre éditorial, la réécriture repose sur une compréhension du contenu et sur des choix assumés concernant ce qui peut être modifié, supprimé ou explicité. À l'inverse, dans le cas de IA, la réécriture s'appuie sur des traitements algorithmiques qui produisent des reformulations linguistiquement plausibles, sans reposer sur une compréhension intentionnelle comparable à celle d'un humain (Julia, 2019).

Il semble intéressant de noter que le geste reste le même en apparence — transformer un texte — mais celui qui le fait n'est plus forcément un humain, encore moins un enseignant avec une intention pédagogique claire.

Dans le cadre éditorial classique, la simplification ne se fait pas au hasard. Chaque mot peut être discuté, chaque phrase peut être ajustée en fonction d'un public précis.

Il y a une forme de responsabilité derrière cette réécriture. On sent qu'il y a quelqu'un qui assume les pertes, qui décide ce qui peut être supprimé sans déformer l'essentiel. C'est exactement ce qui décrit Yves Chevallard (1982, 1985, 2002) quand il parle de « transposition didactique » : un savoir est retravaillé pour devenir apprenable, et ce travail suppose une compréhension fine des enjeux. Rien n'est automatique dans ce processus. Il y a une intention d'enseigner, presque une éthique de la simplification. Quand on passe à l'IA, ce cadre disparaît en grande partie. Le texte n'est plus interprété au sens humain du terme ; il est traité. Cette bascule semble assez brutale. On ne change pas seulement d'outil, on change de logique.

Karsenti (2018, p. 117) le formule d'une manière intéressante : « L'IA peut transformer le travail de l'enseignant en l'amenant, par exemple à jouer un rôle de facilitateur au lieu de transmetteur de contenu ». Cette idée peut paraître séduisante au départ. Pourtant, en y regardant de plus près, il y a quelque chose qui coince. Faciliter quoi exactement, si le texte de départ a déjà été transformé sans reposer sur une compréhension intentionnelle du sens ? On délègue une partie du travail sans être sûr de ce qui a été réellement fait.

Ce décalage devient encore plus visible quand on s'intéresse au rapport au sens. Un éditeur humain sait qu'il simplifie au prix de certaines nuances. Il peut revenir en arrière, hésiter, corriger, reformuler autrement. Il garde une forme de contrôle sur ce qu'il perd. L'algorithme, lui, ne perd rien au sens où il ne sait même pas ce qu'il y avait à perdre. Il privilégie des phrases correctes, fluides, plausibles. Ça sonne juste, mais ça ne garantit pas que ça pense juste.

Julia (2019) le rappelle clairement : les systèmes actuels ne possèdent pas une intelligence comparable à celle des humains, ils reposent sur des traitements de données. « Comprendre » un texte n'est donc pas vraiment ce qu'ils font. À partir de là, un doute s'installe assez vite. Une simplification peut devenir trop explicative, ou au contraire glisser légèrement vers autre chose sans que personne ne s'en rende compte immédiatement. Ce n'est pas spectaculaire, mais c'est précisément ce glissement lié au processus de simplification qui rend la situation problématique.

On peut avoir l'impression que la différence la plus importante ne tient pas aux performances techniques. Elle se situe plutôt dans la posture. D'un côté, quelqu'un qui simplifie pour transmettre quelque chose à quelqu'un d'autre. De l'autre, un système qui produit une version acceptable sans intention de transmission. Cette nuance, d'abord discrète mais essentielle, s'impose progressivement comme un point central de l'analyse menée dans notre recherche notamment pour comprendre et évaluer les implications pédagogiques des différentes formes de simplification.

La simplification automatique ne remplace donc pas vraiment la médiation humaine. Elle la déplace, elle la fragilise aussi d'une certaine manière. Le travail ne disparaît pas, il change de forme, et ça oblige à rester attentif à ce qui circule réellement entre le texte d'origine et celui qui le lit ensuite.

1.4.2. Fonctionnement des modèles de langage

Après avoir distingué la simplification éditoriale de celle produite par l'IA, une évolution du questionnement s'impose progressivement : il ne porte plus uniquement sur le texte simplifié, mais sur le système qui produit cette simplification. Wilkinson et al. (2024, p. 105) inscrivent ChatGPT dans cette dynamique en rappelant que ChatGPT, en tant que modèle de langage de grande taille (Large Language Model - LLM), se situe à l'avant-garde des avancées en IA, une formulation qui met l'accent sur son positionnement technologique, mais qui ne dit encore rien de ce qui se passe réellement à l'intérieur du système.

Morin et Raynault (2023, p. 2) réduisent immédiatement cette image en parlant d'un « robot conversationnel » construit sur l'« apprentissage supervisé pour comprendre et traiter le langage naturel », et cette idée de compréhension reste flottante, comme si elle glissait entre deux sens sans jamais vraiment se stabiliser. Une machine qui comprend, l'expression dérange un peu, surtout quand on la replace dans ce que Morin et son collègue (2023, p. 4) ajoutent ailleurs, à savoir un fonctionnement nourri par un « vaste ensemble de données textuelles disponibles sur Internet », comme si tout reposait moins sur une intelligence que sur une accumulation massive de traces déjà produites.

Le fonctionnement devient plus lisible lorsqu'on quitte les formulations générales. Ji et al. (2026, p. 3) le disent sans détour en indiquant : « At their core, language models are models that assign probabilities to upcoming words, or sequences of words. » Autrement dit, ces modèles reposent sur une estimation probabiliste des unités linguistiques susceptibles de suivre une séquence donnée. Cette formulation semble d'abord offrir une certaine transparence du mécanisme du point de vue du lecteur non spécialiste. Toutefois, cette transparence relève davantage d'une impression de lecture que d'une explication théorique complète du fonctionnement interne. Une suite de tokens apparaît, puis quelque chose se calcule, une sorte de tri silencieux où chaque mot possible reçoit un score, une probabilité, et celui qui gagne devient le mot suivant. Rien d'intuitif ici, plutôt une mécanique répétée, presque obstinée, qui avance mot après mot sans jamais savoir où elle va exactement. Le texte final donne une impression de continuité, parfois même de logique interne, mais cette continuité semble venir après le calcul, comme une illusion produite après le traitement.

Vaishya et al. (2023) insistent sur cette impression troublante de réalisme, en évoquant des productions parfois indifférenciables de l'écriture humaine, ce qui pousse à croire qu'il y aurait une proximité plus grande qu'elle ne l'est réellement. Pourtant, Bender et al. (2021, p. 617) cassent cette idée avec une image qui reste difficile à contourner, celle du « perroquet stochastique » (« stochastic parrot »), un système qui recolle des fragments linguistiques sans jamais renvoyer au sens. L'expression est presque violente, elle oblige à reconsidérer ce qui semblait fluide. Blender et al. (2021) vont encore plus loin en affirmant que le modèle n'a pas accès au sens, et cette absence ne ressemble pas à un manque temporaire mais plutôt à une limite structurelle, quelque chose qui ne se résout pas avec plus de données ou plus de puissance de calcul.

Cette distance entre apparence de compréhension et fonctionnement probabiliste du modèle devient particulièrement problématique lorsqu'on envisage des usages pédagogiques de simplification en FLE. Dans ce cadre, adapter un texte au niveau A2 en FLE ne se réduit pas à une opération mécanique de simplification lexicale ou syntaxique. Il apparaît ainsi que la simplification ne peut être réduite à une opération

technique, mais engage une réflexion plus large sur les conditions de productions du sens, car il y a derrière cela quelque chose de plus subtil, presque une manière d'anticiper la rencontre entre le lecteur et le texte, d'imaginer les zones de résistance possible, les moments où le sens peut devenir moins accessible. Cette dimension, souvent implicite, donne au travail de simplification une vraie profondeur cognitive et pédagogique.

Face aux modèles de langage, on observe une autre logique de production, centrée sur la cohérence linguistique et la fluidité des formes. Le résultat peut être très convaincant, parfois même proche de ce qu'on attend dans un texte simplifié. Mais il reste difficile d'identifier une intention de médiation clairement stabilisée, du moins telle qu'elle apparaît dans les dispositifs d'écriture pédagogique classiques, puisqu'elle ne repose pas sur une situation d'enseignement concrète ni sur un lecteur identifié.

Cela ouvre plutôt un espace d'interrogation sur les formes que peut prendre la simplification lorsqu'elle est déconnectée d'un contexte pédagogique explicite. Cela nous invite alors à réfléchir à la nature même de la simplification : entre opération linguistique, geste pédagogique et construction d'une relation au lecteur. Et c'est peut-être dans cet espace intermédiaire que se joue aujourd'hui une partie importante de la réflexion en didactique du FLE.

1.4.3. Le prompt comme outil de contrôle

Piloter une IA pour simplifier un texte ne relève pas d'un simple geste technique, et c'est peut-être là que le malaise commence. L'idée d'un déclenchement quasi automatique ne tient pas longtemps dès que l'on tente concrètement de formuler une consigne un peu précise, comme si la machine obligeait en retour à clarifier ce que simplifier veut vraiment dire. Cette évolution reste discrète, mais elle redéfinit légèrement le rôle de l'humain. Tôn (2025, p. 27) le formule de manière directe : « người giáo viên cần có các kỹ năng sử dụng thành thạo các công cụ đó » (L'enseignant doit posséder les compétences nécessaires pour utiliser couramment ces outils – Notre traduction). Au-delà de cette exigence technique émerge une forme de dépendance nouvelle à l'outil, une négociation permanente avec lui. Le prompt prend alors place dans cet espace d'incertitude, ni totalement maîtrise humaine ni totalement automatisé.

Ce qui trouble davantage, c'est la manière dont la machine semble répondre, corriger, ajuster, comme si elle entrait dans un dialogue, alors qu'il ne s'agit que d'un enchaînement de calculs. Ganascia (2022, p. 2) parle simplement de « d'automatiser certaines tâches », formule minimale, presque rassurante, mais qui masque mal l'ampleur du glissement : la réécriture, le choix du lexique, l'allègement syntaxique passent progressivement du côté de la machine, sans jamais y appartenir complètement. Les usages répétés donnent parfois l'impression d'une amélioration ; Wilkinson et al. (2024) mettent en évidence que plus l'outil est sollicité, plus les réponses tendent à se préciser et à s'adapter aux attentes de l'utilisateur. Cette accumulation produit des réponses plus ajustées, mais elle installe aussi une impression de précision, difficile à cerner vraiment. Dans un contexte pédagogique, cela peut donner l'impression qu'un texte est adapté, alors qu'il ne correspond pas forcément aux besoins réels d'un groupe d'apprenants. On peut penser à une consigne qui semble claire sur le papier, mais qui laisse certains étudiants silencieux, comme s'ils ne trouvaient pas leur place dans la version proposée. La simplification cesse alors d'être un résultat stabilisé, elle devient une série d'essais, de retours, de reformulations, un processus qui n'en finit pas de se corriger lui-même.

Cette instabilité n'est pourtant pas entièrement nouvelle. Sous une autre forme, elle prolonge des tensions déjà présentes dans la simplification pédagogique. Cette logique n'est pas sans rappeler, de manière un peu décalée, ce qui décrivait Delbart (1994) bien avant l'arrivée de ces outils. Elle évoquait des contraintes très concrètes, imposées à l'écriture simplifiée:

[...] les textes proposés dans ces versions simplifiées recherchent l'expression qui se rapproche le plus de la langue parlée : l'allègement et la simplification syntaxiques sont, avec la sélection du vocabulaire, les impératifs imposés aux auteurs.

(Delbart, 1994, p. 134)

Ce qui change aujourd'hui ne tient pas tant à la nature de ces impératifs qu'à leur lieu d'inscription. L'auteur disparaît partiellement derrière celui qui rédige le prompt, une figure moins visible, moins reconnue, mais qui doit pourtant anticiper, contraindre, orienter. Cette translation n'a rien anodin. Konti et Varsamidou (2025, p. 80) rappelle que « L'IA générative et plus précisément ChatGPT est venu pour faciliter le travail des enseignants [...] et non pas pour les remplacer. » La phrase semble poser une limite claire, mais dans la pratique, la frontière se brouille, surtout lorsque l'outil commence à produire des textes qui paraissent déjà acceptables. À ce moment-là, il devient difficile de savoir s'il faut encore intervenir, ou simplement laisser le texte tel qu'il apparaît.

Un point de friction apparaît dès que l'on sort de cette interaction écran-texte. La relation pédagogique, elle, ne se laisse pas réduire à une suite de consignes. Il suffit d'un décalage minime — un étudiant fatigué, une incompréhension silencieuse, une référence culturelle qui ne passe pas, parfois même un regard qui se détourne, ou une hésitation qui s'installe sans être formulée — pour que tout le dispositif vacille. Tồn (2025, p. 27) le souligne encore : « Việc duy trì mối quan hệ giữa giáo viên và học sinh trong bối cảnh công nghệ ngày càng phát triển cũng là một thách thức lớn » (Le maintien de la relation entre l'enseignant et l'élève dans un contexte technologique en pleine expansion est également un défi majeur – Notre traduction). La machine,

aussi réactive soit-elle, reste extérieure à cette épaisseur du réel, elle ne capte ni les hésitations ni les micro-ajustements qui font le travail quotidien de l'enseignant. Cette limite devient encore plus visible lorsque l'on considère la diversité des situations d'enseignement. Defays et Deltour (2003) décrivent une complexité difficile à contourner :

Enseigner les FLES et donner des conseils à ceux qui le font ne sont pas choses aisées tant les circonstances de cet enseignement peuvent être différentes en fonction de l'endroit où il a lieu (le pays, l'institution...), du profil des apprenants (leur âge, leur langue maternelle, leur niveau linguistique, leur degré d'instruction, leur motivation...), des objectifs poursuivis (scolaires, professionnels, sociaux, culturels...), etc.

(Defays et Deltour, 2003, p. 9)

Face à cette multiplicité, le prompt apparaît presque comme une simplification de la simplification elle-même. Il impose des critères, fixe un niveau, ajuste un ton, mais il reste en surface. La tentation est grande de croire qu'une consigne bien formulée suffit, qu'il existe quelque part une version "correcte" d'un texte simplifié. L'expérience contredit assez vite cette idée. Un même passage retravaillé pour deux groupes différents produit des effets opposés, et aucune ligne de commande ne permet d'anticiper cela complètement.

Il reste alors une impression un peu instable. Simplifier avec une IA fonctionne, souvent, mais jamais de manière totalement satisfaisante. Le prompt devrait agir comme un outil de contrôle, il cadre, oriente, tente de réduire certaines dérives, sans jamais atteindre une forme de maîtrise pleine. Quelque chose échappe, systématiquement. C'est dans cet espace d'ajustement non entièrement calculable que se maintient la place de l'enseignant, qui demeure un médiateur essentiel entre les productions de l'IA et les besoins spécifiques des apprenants.

Au terme de ce chapitre, les différentes notions présentées permettent de mieux comprendre le cadre théorique de cette recherche. Si chacune a été abordée

séparément, elles sont en réalité étroitement liées. La figure suivante résume les principales relations entre les concepts présentés dans ce chapitre.

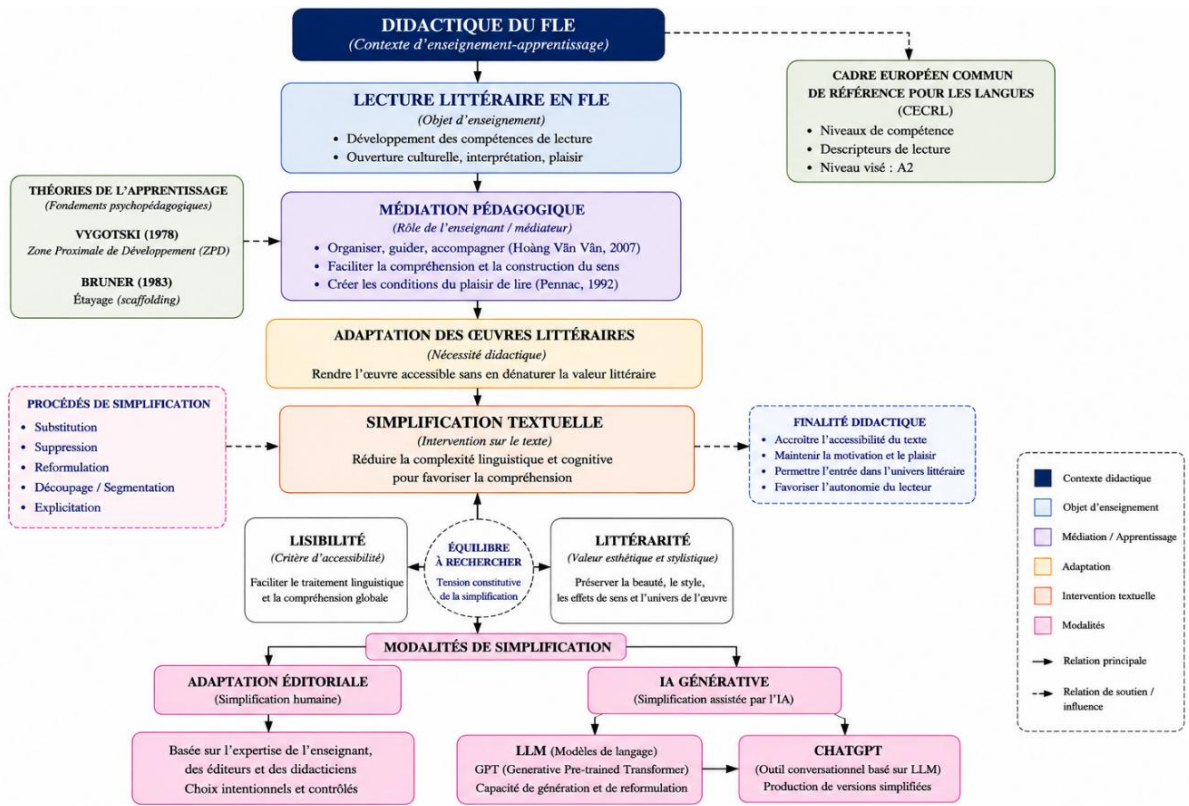


Figure 1. Articulation du cadre théorique de la recherche

CHAPITRE 2 : MÉTHODOLOGIE

2.1. Introduction du chapitre

Le passage de la théorie à la pratique constitue une étape indispensable dans toute recherche scientifique, d'autant plus lorsqu'il s'agit d'évaluer les capacités d'une technologie en constante évolution comme l'IA. Après avoir présenté les bases théoriques de la simplification textuelle et les caractéristiques du style zolien dans le chapitre précédent, nous consacrons le présent chapitre à l'application concrète de notre démarche comparative à plusieurs versions d'un même extrait tiré d'*Au Bonheur des Dames* d'Émile Zola.

Pour ce faire, nous avons cherché à construire un protocole d'analyse destiné à observer précisément les modifications apportées au texte original tout en tenant compte de la richesse du style naturaliste. Ce choix méthodologique ne correspond pas uniquement à une succession d'étapes techniques : il nous permet d'examiner, à partir d'exemples précis, la manière dont la simplification influe sur certains traits stylistiques, tels que la description, le lexique et l'organisation des phrases.

2.1.1. Justification d'une approche hybride et contrastive

L'élaboration d'un protocole méthodologique rigoureux s'est imposée au regard de la densité et de la complexité du texte d'Émile Zola. En observant les passages du premier chapitre d'*Au Bonheur des Dames*, nous avons constaté que certaines phrases, malgré leur longueur, restaient très accessibles à un lecteur averti, tandis que d'autres, même raccourcies, perdaient toute clarté à cause de termes lexicaux comme « soies » ou « dentelles ». Il nous est alors apparu évident que l'étude isolée du texte ne suffirait pas : il fallait comparer plusieurs versions pour discerner les transformations effectives. Nous avons donc placé en regard trois objets distincts : le texte source, l'adaptation de CLE Internationale niveau A2, et une version générée par l'IA (ChatGPT). Cette mise en parallèle permet vraisemblablement de dégager les mécanismes effectifs de la simplification. Sans confrontation directe, il nous paraîtrait délicat de distinguer ce qui relève d'une fidélité littéraire minimale de ce qui correspond à une réduction lexicale ou syntaxique.

Nous avons parallèlement choisi d'articuler une démarche quantitative et une approche qualitative. Ce choix s'explique par les limites des outils automatiques. Les indices de lisibilité, comme le score de Flesch, fournissent des résultats chiffrés, mais ils rendent difficilement compte des nuances de sens ou des effets stylistiques propres au naturalisme. À eux seuls, ils ne permettent donc pas d'apprécier la manière dont un lecteur de niveau A2 comprend réellement le texte. Nous avons ainsi observé qu'un paragraphe consacré aux étals du magasin pouvait obtenir un bon score de lisibilité tout en restant difficile à comprendre en raison du vocabulaire spécifique des marchandises. À l'inverse, certaines phrases longues, enrichies de dialogues ou de descriptions concrètes, restaient étonnamment compréhensibles. Ces constats nous ont convaincue que seule la combinaison de chiffres et d'interprétation humaine permettait une lecture nuancée du phénomène. C'est précisément sur cette complémentarité que repose l'objectivation des résultats.

L'intégration du regard des enseignants de FLE constitue un troisième axe de notre dispositif. Dans cette recherche, nous considérons que la lisibilité ne peut pas être évaluée à partir des seuls indices statistiques. Elle dépend aussi du contexte d'enseignement et des caractéristiques du public auquel le texte est destiné, en l'occurrence des apprenants de niveau A2. C'est pourquoi nous avons fait appel à des enseignants de FLE ayant une expérience de terrain. Leurs évaluations reposent donc sur une pratique pédagogique réelle. Par ailleurs, afin de réduire les risques de biais, les textes ont été présentés dans le cadre d'une évaluation en aveugle. Ce dispositif permettait d'éviter que les jugements soient influencés par l'origine supposée des textes, qu'ils soient attribués à un éditeur ou à une IA. L'anonymisation des versions contribue à nuancer ces préjugés.

Il nous semble ainsi que la triple articulation — comparaison textuelle, mesures objectives, appréciation humaine — n'est pas un luxe, mais la seule manière de rendre compte fidèlement de l'impact de la simplification sur la lisibilité et sur la littérarité. Chaque méthode éclaire un aspect différent du phénomène étudié. Et c'est en confrontant nos propres observations aux résultats chiffrés que nous avons commencé

à comprendre, de manière très concrète, comment Zola est transformé sous la plume humaine et sous celle de l'IA.

2.1.2. Présentation générale du dispositif méthodologique

Dans cette sous-section, nous décrivons l'architecture générale du protocole de recherche. Elle présente la manière dont les différentes étapes sont organisées et reliées entre elles. Le schéma ci-dessous synthétise cette organisation et illustre le flux de traitement des données.

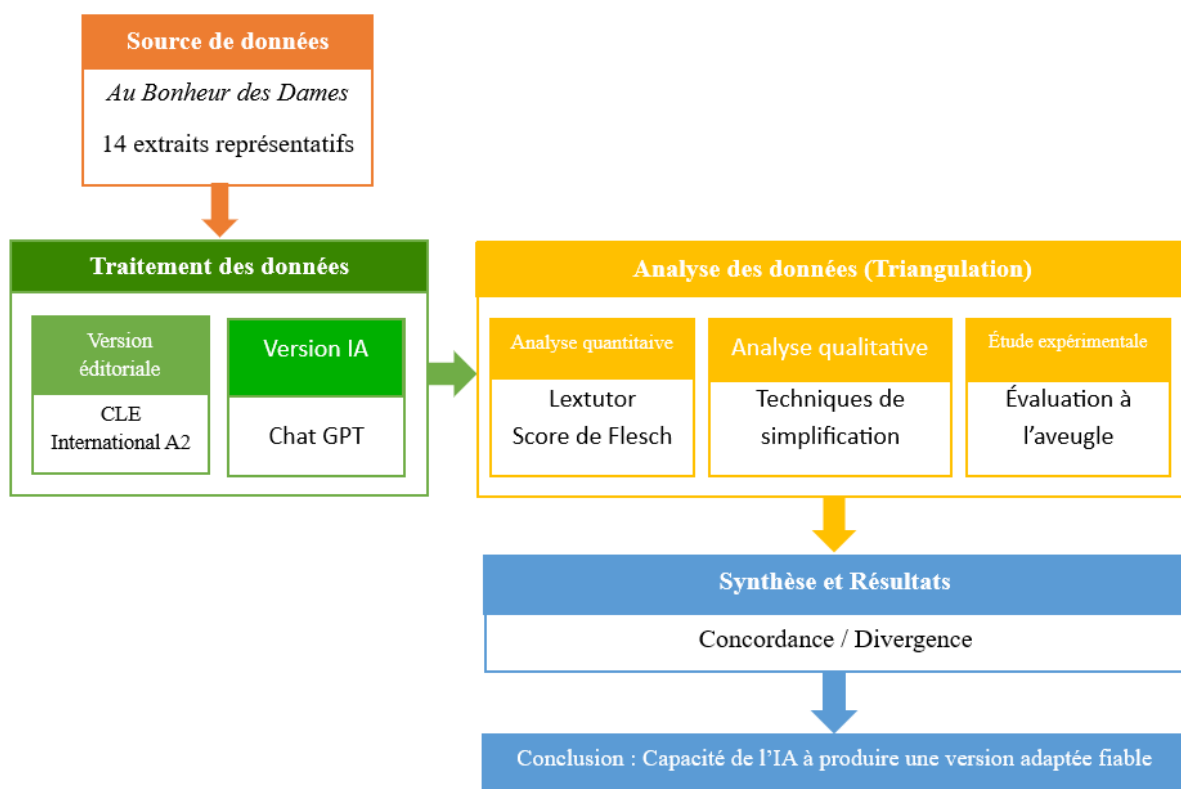


Figure 2. Schéma de la méthodologie de recherche

Comme le montre la Figure 1, le protocole de recherche suit un flux organisé : chaque phase prend place à l'issue d'une autre et s'y articule étroitement. La première étape a consisté à constituer le corpus de travail. Nous avons retenu quatorze extraits de *Au Bonheur des Dames* d'Émile Zola, qui servent de point de départ à l'ensemble du protocole. Tous les extraits sont conservés dans leur version originale afin que les comparaisons reposent sur une même base, tout en couvrant des situations narratives et des caractéristiques linguistiques variées.

Chaque extrait a ensuite été décliné en deux versions simplifiées : une adaptation éditoriale et une version produite par ChatGPT. Les différentes versions ont été regroupées de manière à suivre facilement le parcours de chaque extrait tout au long de la recherche. Un tableau récapitulatif permet d'identifier chaque texte à chaque étape du protocole. Il sert avant tout à assurer la cohérence de l'ensemble et à préparer les données pour les analyses ultérieures, sans porter d'appréciation sur la qualité des différentes versions.

Après la stabilisation des différentes versions, l'analyse s'organise autour de trois dimensions. La première concerne les données quantitatives, qui permettent d'obtenir des indications mesurables sur la lisibilité des textes à partir du score de Flesch et d'autres indices automatisés. La deuxième repose sur une comparaison qualitative des versions éditoriales et des versions générées par ChatGPT, afin d'identifier les principales transformations au niveau des structures et des formulations. Enfin, une troisième partie s'appuie sur l'évaluation en aveugle menée auprès d'enseignants de FLE, dont les observations viennent compléter les analyses précédentes. Nous avons décidé de relier chaque dimension au tableau central afin de suivre la progression et la correspondance des résultats, tout en laissant des liens visibles entre les axes pour que le protocole reste flexible et réutilisable. Le détail des indicateurs et des critères propres à ces dimensions est exposé progressivement dans les sections suivantes : la grille d'analyse multidimensionnelle (section 2.3) pour les volets quantitatifs et qualitatifs, ainsi que le cadre de l'étude expérimentale (section 2.4.4) pour l'évaluation humaine.

Enfin, les résultats obtenus dans les différentes phases d'analyse sont rassemblés afin de préparer leur mise en relation. Les données sont classées selon leur origine (quantitative, qualitative ou évaluative) et réparties dans des catégories distinctes pour conserver une lecture claire du protocole. À cette étape, il ne s'agit pas encore d'évaluer la qualité des textes, mais simplement d'organiser les informations recueillies afin de pouvoir les comparer et les interpréter dans la phase de synthèse.

Dans l'ensemble, le protocole adopte une logique de flux clair : le corpus est traité, les versions circulent vers l'analyse selon trois axes, et les résultats sont organisés

pour une synthèse finale. Les relations entre étapes sont présentées explicitement dans le schéma, permettant de suivre le cheminement des données de manière transparente, tout en conservant un regard neutre sur le contenu lui-même.

2.2. Constitution et description du corpus

2.2.1. Texte source : *Au Bonheur des Dames* d'Émile Zola

Publié en 1883, *Au Bonheur des Dames* appartient au cycle des *Rougon-Macquart*, ensemble romanesque dans lequel Zola (1873) se propose d'étudier, selon une perspective naturaliste, « l'histoire naturelle et sociale d'une famille sous le Second Empire ». Le roman s'inscrit dans un projet littéraire plus large, fondé sur l'observation des milieux, la mise en relation des trajectoires individuelles avec des forces économiques et sociales, ainsi que sur une attention constante portée aux mécanismes collectifs. Dans cette série, *Au Bonheur des Dames* occupe une place spécifique : il est consacré à la naissance et au développement du grand magasin parisien, inspiré notamment des établissements réels apparus dans la seconde moitié du XIX^e siècle. Le récit met en scène l'essor d'un commerce moderne fondé sur l'abondance des marchandises, la scénographie de la vente et la transformation des pratiques de consommation.

Le grand magasin n'y constitue pas un simple décor. Il organise l'ensemble du dispositif narratif et agit comme un véritable pôle d'attraction autour duquel se structurent les relations entre les personnages. Plusieurs critiques ont souligné cette fonction centrale de l'espace commercial. Dans ses analyses consacrées à Zola, Mitterand (1990) montre que le magasin est construit comme une machine sociale, régie par une logique d'expansion continue et par une scénographie destinée à capter le regard du client. Cette conception se traduit dans le roman par une écriture qui multiplie les descriptions d'ensemble, les effets de perspective et les mouvements de foule, donnant au lieu une dimension presque organique. Le magasin devient à la fois un espace économique et un symbole de la modernité urbaine, ce qui explique la place importante accordée à sa représentation matérielle.

Sur le plan linguistique, le texte se distingue par une forte densité descriptive. Comme l'analyse Philippe Hamon (cité par Saint-Gérard, 1993), la description zolienne ne

constitue pas un simple ornement, mais s'apparente à un « exercice de la nomination » qui vise à « épuiser la réalité du monde » par une « ambition d'exhaustivité ». Cette volonté se manifeste par une syntaxe particulièrement saturée : les phrases sont souvent longues, structurées autour de plusieurs propositions subordonnées et de nombreuses expansions du nom. Le lexique mobilisé est particulièrement riche dans les passages consacrés aux marchandises : Zola recourt à des termes techniques pour désigner les tissus et les étoffes, tels que « faille » ou « guipure », qui renvoient à des réalités matérielles précises mais peu familières pour un lecteur non spécialiste. À cette précision lexicale s'ajoutent des procédés d'énumération et d'accumulation visant à restituer l'impression d'abondance propre au grand magasin, mais qui, selon les modèles de Giasson (1996), augmentent considérablement la charge cognitive (Sweller et al., 2007) nécessaire à la compréhension du sens.

Cette complexité linguistique se double d'une hétérogénéité discursive marquée. Le roman alterne en effet des séquences narratives, centrées sur l'évolution des personnages, et de longs passages descriptifs consacrés aux espaces, aux objets et aux flux de clients. Pour un apprenant de niveau A2, cette alternance représente un obstacle didactique majeur, car la présence d'un lexique spécialisé, notamment dans le domaine du textile et du commerce, suppose des connaissances qui dépassent largement le « répertoire lexical » de base (environ 1000 à 1200 mots) habituellement maîtrisé à ce niveau selon les descripteurs du CECRL (2001). Comme le souligne Cuq (2003), certains obstacles à la compréhension viennent de termes techniques peu transparents et de groupes nominaux complexes. Dans le cas d'un passage consacré à l'étalage des étoffes, le lecteur ne doit pas seulement identifier les noms utilisés. Il doit également comprendre les liens entre les différents éléments de la phrase et reconstruire les relations de sens qui organisent la description. Sans accompagnement ou travail préalable de simplification, ce type de passage peut donc représenter une difficulté importante pour les apprenants.

Ces éléments font d'*Au Bonheur des Dames* un corpus particulièrement intéressant pour étudier les mécanismes de simplification linguistique dans une perspective

didactique. Le roman met en effet en évidence un écart important entre la complexité de son écriture originale et les contraintes de lisibilité liées à un public débutant ou élémentaire. À partir d'un même texte, il devient possible d'observer différentes transformations : la réduction de certaines constructions syntaxiques longues, la modification d'un vocabulaire spécialisé ou encore la réorganisation de passages descriptifs. Ces transformations permettent d'étudier comment un texte littéraire peut être rendu plus accessible sans que l'ensemble de ses informations soit nécessairement supprimé.

Dans cette recherche, le roman est envisagé principalement comme un texte-source. Les extraits sélectionnés constituent le matériau initial à partir duquel sont produites les différentes versions étudiées. Ils sont prélevés dans l'œuvre originale, sans adaptation préalable, afin de conserver un point de comparaison commun entre les textes. Le roman n'est donc pas analysé ici uniquement pour sa valeur esthétique, mais comme un objet linguistique permettant d'examiner différents procédés de reformulation et de simplification.

Le choix de conserver un même texte de départ permet ainsi de comparer les transformations réalisées dans les différentes versions. Qu'elles proviennent d'une adaptation éditoriale ou d'une production par intelligence artificielle, toutes les reformulations peuvent être rapportées au même support initial. *Au Bonheur des Dames* occupe donc une place centrale dans le corpus : il constitue le texte de référence à partir duquel les analyses quantitatives et qualitatives sont menées.

2.2.2. Critères de sélection des quatorze extraits

La constitution du sous-corpus repose sur une sélection raisonnée de quatorze extraits tirés de *Au Bonheur des Dames*. Nous ne procédons ni par découpage aléatoire, ni par simple échantillonnage quantitatif. La sélection résulte d'un processus graduel, visant deux objectifs complémentaires : assurer la représentativité de l'écriture zolienne et permettre une analyse comparative entre différentes formes de simplification. Pour faciliter la lecture de notre analyse, nous présentons ci-dessous un tableau récapitulatif des quatorze extraits (identifiés par les codes de E1 à E14), qui

constituent le corpus de recherche. Le texte intégral, comprenant les versions originale et simplifiée, figure en *Annexe 2*.

Tableau 1. *Inventaire et contenu narratif des quatorze extraits du corpus*

Code	Titre de l'extrait	Localisation (Chapitre / Page)	Longueur (Mots)	Aspects narratifs et thématiques saillants
E1	Installation de Denise	Ch. 4 (pp. 141-142)	182	Arrivée de Denise à Paris et passage symbolique de sa robe de laine (origine provinciale) à l'uniforme de soie (intégration au magasin).
E2	Stratégies de séduction commerciale	Ch. 9 (p. 325)	155	Mise en place d'innovations technologiques et stratégie de séduction des clientes par le biais de l'enfant.
E3	Progrès humanitaires et cité du travail	Ch. 12 (pp. 476-477)	136	Modernisation sociale du magasin offrant une autonomie complète aux employés (éducation, culture, soins médicaux et logement).
E4	Apothéose du blanc	Ch. 14 (pp. 527-529)	363	Spectacle visuel spectaculaire de tous les articles de couleur blanche.
E5	Portrait de Mouret	Ch. 2 (pp. 69-70)	143	Portrait physique d'Octave Mouret et naissance d'un trouble émotionnel complexe chez Denise.
E6	Évolution de la fratrie Baudu	Ch. 14 (pp. 538-539)	115	Portrait de la croissance corporelle spectaculaire des frères Baudu, devenus des « grands gaillards » dominant Denise par la taille, tandis que celle-ci réaffirme son autorité

				maternelle et son dévouement protecteur.
E7	Frénésie de la foule dévorante	Ch. 9 (pp. 331-333)	315	Scène où la foule des clientes devient une force organique et dévorante, mue par la passion de l'achat.
E8	Détresse morale et fatigue physique	Ch. 4 (p. 180)	171	Souffrance physique et détresse morale de Denise face à la solitude et à la dureté de sa première journée de travail.
E9	Vanité de la fortune et vide intérieur	Ch. 12 (pp. 481-482)	313	Crise existentielle de Mouret qui, malgré sa fortune et l'expansion du magasin, souffre du vide affectif lié à Denise.
E10	Prise de conscience amoureuse	Ch. 10 (p. 370)	257	Lecture de la lettre de Mouret ; Denise réalise la nature de son amour et les enjeux de sa réputation au sein du magasin.
E11	Confrontation avec Mme Desforges	Ch. 11 (pp. 431-432)	226	Humiliation de Denise par Henriette Desforges et intervention protectrice de Mouret, révélant sa tendresse croissante.
E12	Licenciement et dignité	Ch. 6 (pp. 254-255)	214	Licenciement brutal de Denise par Bourdoncle ; elle quitte le magasin avec une dignité silencieuse malgré sa détresse.
E13	Affirmation de soi et refus du luxe	Ch. 10 (pp. 403-405)	371	Refus catégorique de Denise de l'offre de fortune de Mouret,

				affirmant sa liberté et sa dignité par le travail.
E14	Dénouement et aveu réciproque	Ch. 14 (pp. 572- 573)	323	Aveu réciproque des sentiments dans un moment de désespoir mutuel, marquant la résolution finale de l'intrigue.

La sélection du corpus ne repose pas sur un seul critère. Trois dimensions ont été prises en compte : les caractéristiques textuelles des extraits, leurs propriétés linguistiques ainsi que leur intérêt pour une analyse didactique. Les différents choix effectués sont présentés selon ces trois niveaux.

Le premier élément examiné concerne la nature même des passages retenus. La longueur des extraits a notamment constitué un point d'attention important. La plupart des textes sélectionnés comptent entre environ 120 et 370 mots, pour une moyenne d'environ 245 mots. Cette variation permet de travailler sur des unités suffisamment riches pour observer les modifications apportées par la simplification, tout en évitant des passages trop étendus qui auraient rendu les comparaisons plus difficiles. Certains écarts existent néanmoins dans le corpus : l'extrait 6, avec ses 115 mots, représente l'un des passages les plus courts, tandis que l'extrait 13 atteint 371 mots. Ces différences restent limitées et ne remettent pas en cause la cohérence générale de l'ensemble.

La diversité des situations narratives a également orienté le choix des extraits. Le corpus ne se limite pas aux descriptions, même si celles-ci occupent une place importante dans l'écriture naturaliste de Zola. Certains passages retenus mettent en scène l'organisation du magasin, comme dans l'extrait 2, ou encore la présentation du blanc dans l'extrait 4. D'autres correspondent davantage à des moments narratifs ou dialogués, par exemple le licenciement de Denise (extrait 12) ou son échange conflictuel avec Mme Desforges (extrait 11). La présence de ces différentes formes de discours permet d'observer les effets de la simplification sur des textes aux fonctionnements variés.

Le choix des extraits a ensuite été guidé par leurs caractéristiques linguistiques. Les passages retenus présentent généralement une forte densité lexicale ainsi qu'une organisation syntaxique complexe. On y trouve notamment des phrases longues, des propositions subordonnées imbriquées et des groupes nominaux développés, particulièrement fréquents dans les descriptions d'objets ou d'espaces commerciaux. Le vocabulaire spécialisé constitue également un élément important de cette sélection. L'extrait 7 en offre un exemple représentatif, avec une accumulation de termes et d'expressions décrivant le mouvement des clientes, les bousculades et les différents articles proposés à la vente. La richesse lexicale dépasse largement le répertoire habituellement maîtrisé par des apprenants de niveau A2. À l'inverse, les dialogues très courts ou fortement fragmentés ont été écartés afin de préserver un contraste pertinent entre le texte source et ses versions adaptées. Nous avons ainsi privilégié des segments offrant une « résistance » linguistique suffisante pour permettre une observation tangible des opérations de réduction et de reformulation.

La dimension didactique a également joué un rôle dans le choix des extraits. Nous avons retenu des passages dont la transformation vers un niveau A2 pouvait être observée de manière pertinente, c'est-à-dire des textes où une réduction de la complexité formelle pouvait être envisagée sans entraîner la disparition des informations essentielles du récit. L'objectif n'était donc pas de sélectionner des passages déjà accessibles ou proches d'un texte pédagogique, mais plutôt des extraits suffisamment riches pour permettre d'analyser les effets produits par la simplification. Nous avons également accordé une attention particulière à l'autonomie de chaque extrait. Même s'ils proviennent d'un roman construit sur une intrigue longue, les quatorze passages sélectionnés correspondent à des unités relativement indépendantes : descriptions d'espaces, portraits de personnages comme celui d'Octave Mouret (extrait 5), ou encore scènes marquées par des tensions entre les personnages. Leur compréhension ne nécessite pas une connaissance complète de l'ensemble du roman, ce qui permet de les utiliser comme supports autonomes dans une situation d'enseignement.

Le sous-corpus a enfin été organisé selon une logique thématique et discursive. Les extraits, numérotés de E1 à E14, suivent un ordre établi en fonction des besoins de l'analyse et non selon la progression chronologique du roman. Ce choix permet de mettre en évidence à la fois les caractéristiques linguistiques des passages et les différentes évolutions thématiques présentes dans le corpus. Trois groupes principaux apparaissent ainsi : les extraits E1 à E4 portent principalement sur la description des espaces et sur les stratégies commerciales du grand magasin ; les extraits E5 à E8 sont davantage consacrés aux portraits des personnages ; enfin, les extraits E9 à E14 mettent au premier plan les tensions sociales, les conflits individuels et la résolution amoureuse, avec une dimension plus narrative et discursive. Parmi ceux-ci, les extraits 11 à 14 sont surtout composés de dialogues, ce qui offre un contraste intéressant pour l'analyse des opérations de simplification et de reformulation. Cette structuration garantit que l'analyse ne se limite pas à un seul type de discours, mais qu'elle couvre l'ensemble des registres mobilisés par Zola.

L'ensemble de ces critères permet de constituer un sous-corpus à la fois cohérent et varié. Les quatorze extraits offrent un échantillon représentatif des principales caractéristiques linguistiques et discursives du roman, tout en restant compatibles avec les contraintes d'un protocole comparatif. À partir de ces unités soigneusement sélectionnées, deux types de versions dérivées ont été élaborés pour servir de base à l'analyse : une version éditoriale (VE) et une version générée par l'intelligence artificielle (VIA). Cette étape marque le passage du texte source vers le dispositif comparatif proprement dit.

La figure 2 illustre l'organisation interne du sous-corpus ainsi que les trois séries de critères ayant guidé la sélection des extraits : critères textuels, linguistiques et didactiques. Cette représentation permet de visualiser la répartition des extraits selon les types de séquences romanesques et leur progression analytique.

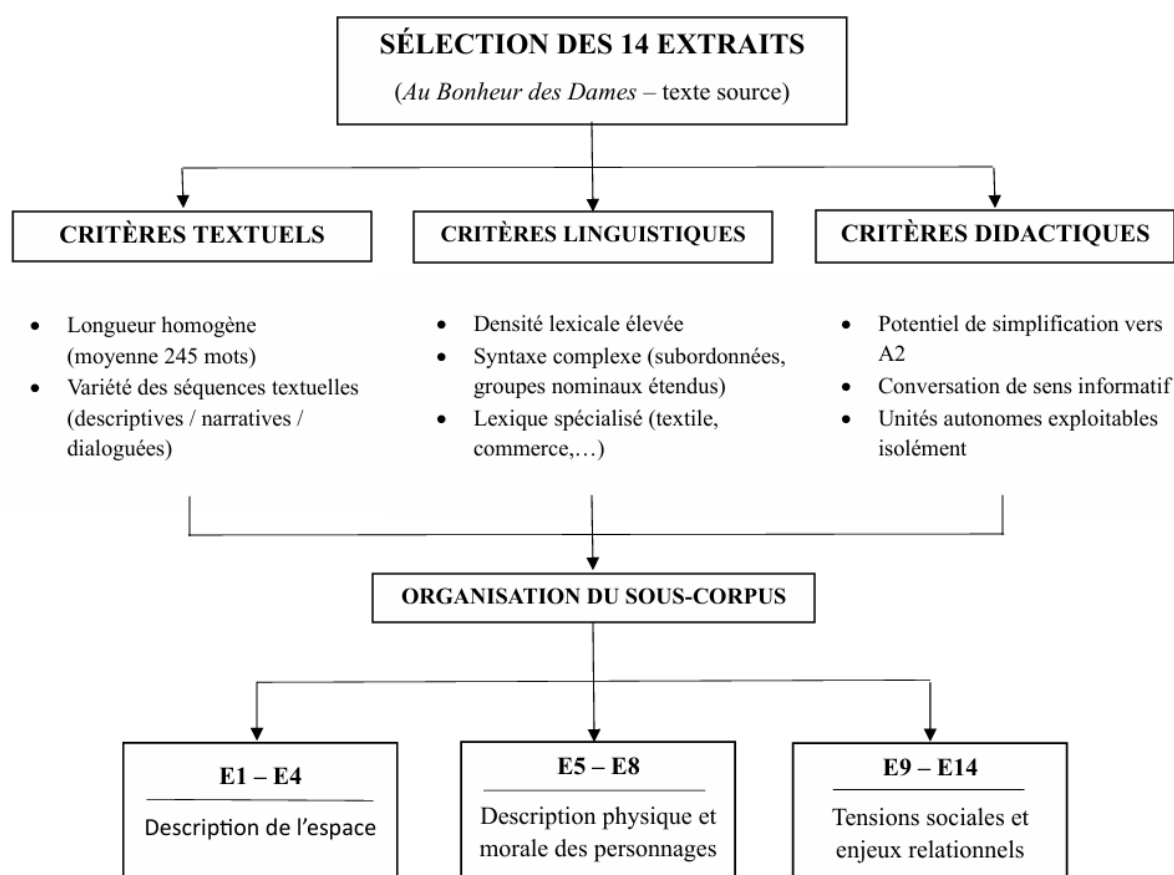


Figure 3. *Organisation et critères de sélection du sous-corpus*

2.2.3. La version éditoriale : l'adaptation de CLE Internationale (Niveau A2)

Le premier volet de notre corpus dérivé est constitué par une adaptation issue du circuit éditorial professionnel. Cette version, que nous désignerons par la suite sous le sigle VE (Version éditoriale), correspond au modèle traditionnel de médiation linguistique proposé aux apprenants de français langue étrangère. Chaque extrait éditorial simplifié correspond à un extrait du texte source initialement sélectionné et est identifié de manière parallèle, de VE1 à VE14, afin de garantir la traçabilité entre les différentes versions d'un même passage.

La version retenue provient de la collection « Lectures CLE en français facile » de la maison d'édition CLE International, adaptée en français facile par Françoise Claustres. Elle est destinée à un public d'apprenants de niveau A2, conformément aux descripteurs du Cadre européen commun de référence pour les langues (CECRL). À la différence du texte original de Zola, cette version n'a pas été conçue comme une

œuvre littéraire autonome. Elle répond avant tout à un objectif pédagogique : proposer aux apprenants un accès plus facile à un texte complexe. Produite dans un cadre éditorial institutionnel, elle résulte d'un travail humain réalisé par des spécialistes de la didactique des langues, chargés d'adapter une matière littéraire aux besoins d'un public d'apprentissage. Dans cette perspective, la VE constitue un exemple représentatif des pratiques de simplification éditoriale utilisées en FLE.

L'analyse des modifications apportées dans la VE fait apparaître plusieurs régularités. La première concerne la structure des phrases. Les périodes longues et les constructions syntaxiques complexes, notamment celles qui reposent sur plusieurs niveaux d'enchâssement, sont souvent remplacées par des phrases plus courtes ou par des structures coordonnées et juxtaposées. Les transformations touchent également le vocabulaire. Les termes rares ou fortement spécialisés sont fréquemment remplacés par des expressions plus générales. C'est notamment le cas du lexique textile présent chez Zola : des mots comme « taffetas » ou « failles » peuvent être remplacés par des termes plus larges tels que « tissu » ou « soies », ou disparaître lorsque leur rôle descriptif est considéré comme moins essentiel. La réorganisation syntaxique vise à alléger la charge cognitive tout en maintenant la cohérence du récit et la progression narrative essentielle.

Il convient toutefois de préciser que, dans certains cas, l'adaptation éditoriale ne conserve qu'une phrase unique issue de l'extrait source. Cette réduction extrême ne résulte pas d'un choix méthodologique de notre part, mais du principe même de la simplification éditoriale, qui privilégie parfois la restitution d'un noyau informatif ou narratif jugé suffisant pour le niveau A2, au détriment du développement descriptif initial. Cette hétérogénéité de longueur entre les extraits originaux et leurs versions éditoriales constitue ainsi une contrainte du corpus, mais elle reflète fidèlement les pratiques effectives de l'édition didactique.

Dans le cadre de notre dispositif méthodologique, la VE occupe une position centrale en tant que point de comparaison humain. Elle sert d'étalon pour l'évaluation des opérations de simplification, représentant ce qu'un éditeur spécialisé et un adaptateur expérimenté considèrent comme acceptable et pertinent pour un public de niveau A2.

En tant que modèle implicite de simplification didactique, elle permet de mener une analyse qualitative comparée entre les choix opérés par l'expert humain et ceux produits par un système automatisé. La confrontation de ces deux types de médiation linguistique vise à déterminer si l'intelligence artificielle reproduit, prolonge ou modifie les logiques éditoriales traditionnellement mises en œuvre dans le domaine du FLE.

2.2.4. La version artificielle : protocole de génération de l'IA

Le second volet du corpus dérivé correspond à une version produite artificiellement par un système d'intelligence artificielle générative. Cette version, que nous désignerons par le sigle VIA (Version par Intelligence Artificielle), repose sur l'utilisation d'un outil unique : Chat GPT, dans son état de fonctionnement au moment de la collecte des données (version GPT-4o, consultation effectuée en février 2026). Le choix de ne recourir qu'à un seul outil vise à limiter la variabilité technologique et à garantir une cohérence minimale entre les productions. Dans le cadre de notre protocole, l'IA est mobilisée non comme source de contenu autonome, mais comme dispositif de reformulation, chargé de transformer un texte littéraire complexe en une version simplifiée destinée à un public d'apprenants de niveau A2.

Avant de constituer le corpus artificiel définitif, nous avons réalisé une première phase d'expérimentation à partir d'un extrait du texte source. Ce test pilote avait pour but d'examiner la précision des consignes données à l'IA et de vérifier si celles-ci permettaient d'obtenir des productions suffisamment stables. Plusieurs formulations du prompt ont ainsi été essayées. Certaines se limitaient à indiquer le niveau linguistique attendu (A2), tandis que d'autres ajoutaient des indications concernant les aspects stylistiques, comme la conservation des métaphores, du ton ou du rythme du texte. Les différentes productions obtenues ont ensuite été confrontées entre elles ainsi qu'au texte original. Cette comparaison a permis d'identifier les situations dans lesquelles l'IA avait tendance soit à simplifier excessivement certains éléments, soit à maintenir un vocabulaire encore trop complexe pour des apprenants de niveau A2. Ce test préalable n'a pas été intégré au corpus analysé. Il a cependant constitué une étape importante pour ajuster le dispositif de génération. Les observations réalisées

ont permis de préciser les consignes finales afin d'obtenir des versions compatibles avec le niveau A2, tout en laissant à l'IA une marge suffisante pour traiter les dimensions linguistiques et stylistiques du texte. Cette étape préparatoire contribue ainsi à rendre la comparaison entre les VIA et la version éditoriale plus cohérente.

Les consignes officielles adressées à l'outil ont été uniformes pour l'ensemble des extraits. Chaque requête précisait le niveau visé (A2 selon le CECRL) et le type de tâche attendu : une simplification linguistique du texte source accessible aux apprenants. Le prompt comprenait également des critères stylistiques adaptés : les métaphores et images stylistiques étaient conservées uniquement si elles demeuraient compréhensibles pour le niveau cible ; sinon, elles étaient reformulées ou simplifiées tout en maintenant le sens et l'effet stylistique. Cette approche vise à évaluer la capacité de l'IA à produire une version simplifiée didactiquement efficace en assurant la lisibilité et la cohérence du texte final. La stabilité de ces consignes constitue une condition essentielle pour maintenir la comparabilité entre les différentes sorties produites.

La procédure de génération a été organisée de manière systématique. Chaque extrait du texte source a été traité séparément et soumis à l'outil dans une session de conversation distincte. Ce fonctionnement vise à éviter toute interférence contextuelle entre les demandes successives et à neutraliser l'effet de mémoire propre aux systèmes conversationnels. À chaque extrait correspond une seule sortie produite par l'IA. Les textes générés ont ensuite été identifiés et classés selon une nomenclature parallèle à celle du corpus source et de la version éditoriale, de VIA1 à VIA14, afin d'assurer la traçabilité entre les différentes versions d'un même passage. Après leur génération, les textes produits par l'IA n'ont fait l'objet d'aucune modification de notre part. Nous avons conservé chaque version telle qu'elle a été obtenue, sans correction linguistique ni réécriture ultérieure. Cette décision permet d'analyser les productions de ChatGPT dans leur forme initiale et d'observer directement les choix effectués par le système lors de la simplification. Ces versions doivent être considérées comme des textes issus d'un processus automatisé, et non comme des exemples à suivre ou des modèles de simplification pour l'enseignement

du FLE. Elles se distinguent ainsi de la version éditoriale, qui résulte d'un travail humain inscrit dans un cadre pédagogique précis. La VIA représente donc avant tout un matériau d'étude permettant d'examiner les possibilités et les limites d'une simplification réalisée par une intelligence artificielle, en comparaison avec une démarche éditoriale humaine.

Dans l'économie générale de l'étude, la version artificielle occupe une place centrale aux côtés de la version éditoriale. Elle constitue l'outil principal pour interroger la capacité d'un système automatisé à produire des reformulations conformes aux exigences d'un niveau A2, tout en permettant un contraste éclairant avec les choix opérés par l'adaptateur humain. La confrontation entre VE et VIA ne cherche pas à hiérarchiser a priori les deux types de production, mais à mettre en évidence les logiques respectives selon lesquelles le texte source est transformé, ainsi que l'efficacité potentielle de l'IA comme dispositif de simplification pour l'enseignement du français langue étrangère.

Les consignes détaillées adressées à l'outil sont présentées en *Annexe 1*.

2.3. Grille d'analyse multidimensionnelle

L'analyse menée dans ce travail repose sur une logique contrastive, fondée sur la confrontation de trois versions d'un même extrait : le texte source, la version éditoriale destinée aux apprenants de niveau A2 et la version générée par ChatGPT. Il apparaît d'emblée que l'examen isolé de ces textes ne permet pas de saisir pleinement les transformations opérées par la simplification. Une lecture intuitive peut suggérer certaines tendances, mais elle demeure fragmentaire. Il nous a donc semblé nécessaire de dépasser une approche impressionniste pour construire un dispositif d'observation plus structuré. Une seule dimension d'analyse, qu'elle soit lexicale ou syntaxique, risquerait de réduire la complexité des phénomènes observés. La simplification ne touche pas uniquement la forme linguistique ; elle affecte également l'organisation du sens, la texture stylistique et, vraisemblablement, la perception de la lisibilité.

Dans cette perspective, nous avons élaboré une grille d'analyse multidimensionnelle destinée à articuler plusieurs niveaux d'observation complémentaires. Si un premier

axe se concentre sur les transformations formelles du lexique et de la syntaxe, le suivant s'attache plutôt aux procédés concrets de réécriture, tels que la suppression, la substitution, le découpage, l'explicitation et la reformulation. Quant au troisième axe, il interroge la relation entre la conservation du sens et les effets de littérarité, souvent fragilisés par la simplification. Un dernier niveau d'analyse introduit enfin une mesure objectivable de la lisibilité à l'aide d'indices quantitatifs, sans prétendre pour autant épuiser la question de la compréhension réelle. Afin de rendre compte de manière synthétique de cette grille, nous en proposons une présentation récapitulative dans le tableau ci-dessous.

Tableau 2. Grille d'analyse multidimensionnelles des textes

Axes d'analyse	Critères et Indicateurs d'analyse	Enjeux théoriques et interprétatifs
Analyse linguistique formelle	- Lexique : Repérage de la nature du vocabulaire (courant, littéraire, concret, spécialisé) et analyse qualitative des substitutions (présence de neutralisation lexicale). - Syntaxe : Observation des simplifications morphosyntaxiques et repérage de la complexité structurelle (subordination, expansions, incises, accumulations).	- Réduction de la charge cognitive. - Simplification du traitement linguistique - Transformation du lexique zolien concret et sensoriel.
Opérations de médiation textuelle	- Catégorisation des procédés : Identification des segments transformés via 5 opérations clés : suppression, substitution, reformulation, explicitation, découpage (segmentation des phrases longues).	- Étayage pédagogique. - Réduction des « blancs ». - Modification du rythme narratif. - Transposition didactique. - Risque de sur-explicitation (réduction excessive des implicites).
Tension entre littérarité et lisibilité	- Style et narration : Analyse qualitative des effets stylistiques, incluant la conservation ou perte des accumulations zoliennes et la transformation des images/descriptions sensorielles. - Sémantique et réception : Observation des	- Horizon d'attente du lecteur. - Préservation ou effacement de la littérarité. - Rapport entre accessibilité et expérience esthétique. - Rapport entre fluidité et résistance du texte.

	glissements sémantiques, vérification du maintien des invariants narratifs, et évaluation du niveau d'explicitation.	- Perte de l'immersion perceptive.
Évaluation objective de la lisibilité	- Mesures statistiques : Calcul de la longueur moyenne des phrases et des mots. - Outils et indices : Application des indices de lisibilité (Flesch, Kandel-Moles) et utilisation de Lextutor pour analyser les pourcentages de répartition lexicale (K1, K2, K3, Off-list).	- Mesure de l'accessibilité linguistique. - Identification du niveau de difficulté. - Recherche d'un équilibre entre simplification et maintien d'un lexique stimulant.

Cette grille ne constitue pas un modèle théorique autonome. Nous l'avons construite à partir de notions et de typologies issues des travaux existants sur la simplification textuelle et l'analyse stylistique, tout en l'adaptant aux contraintes spécifiques de notre corpus. Il ne s'agit donc ni d'un cadre normatif ni d'un outil universel, mais d'un instrument opératoire conçu pour répondre aux besoins précis de cette recherche. Il remplit à la fois une fonction de codage, en permettant de classer les transformations observées, une fonction de comparaison entre les différentes versions d'un même extrait et une fonction interprétative, en éclairant les relations entre simplification linguistique, fidélité sémantique et lisibilité.

L'organisation interne de cette grille repose sur quatre axes correspondant chacun à un niveau d'observation distinct. Cette structuration vise à rendre compte de manière progressive des effets de la simplification, depuis les modifications linguistiques les plus visibles jusqu'aux conséquences plus diffuses sur le sens et la réception du texte.

2.3.1. Analyse linguistique formelle

Comme défini dans le premier axe de notre grille, l'observation linguistique constitue le premier point d'entrée de notre analyse, dans la mesure où la simplification se manifeste d'abord par des transformations visibles dans la matière même du texte. Nous avons choisi d'aborder ces modifications à travers l'examen du lexique et de la syntaxe, afin de décrire les changements formels sans encore leur attribuer une valeur interprétative. À ce stade, il ne s'agit pas d'évaluer la qualité stylistique des versions produites, mais de repérer ce qui est conservé, déplacé ou supprimé.

L'analyse lexicale constitue une première entrée pour observer les effets de la simplification. Dans le texte source, certains passages s'appuient sur un vocabulaire directement lié à l'univers du grand magasin, notamment à travers les termes du commerce et du textile. D'autres mobilisent davantage un lexique abstrait, utilisé pour exprimer les états intérieurs des personnages ou décrire les mouvements collectifs. On trouve également un vocabulaire lié à la narration elle-même, qui participe à l'organisation des actions et à la progression du récit. La comparaison entre les différentes versions permet ainsi d'identifier les catégories lexicales davantage conservées et celles qui font l'objet de modifications ou de suppressions. Cette observation du vocabulaire est complétée par une analyse de la complexité lexicale. Celle-ci n'est pas envisagée comme un jugement de difficulté absolue, mais comme un ensemble de caractéristiques observables dans les textes. Nous examinons notamment les mots peu fréquents, les remplacements de termes spécialisés par des expressions plus courantes, ainsi que la réduction éventuelle des mots comportant plusieurs syllabes. Une attention particulière est également accordée aux dérivations morphologiques, dont la compréhension peut varier selon le niveau des apprenants. Ainsi, le rapport entre « vendre » et « vendeuse » peut être facilement identifié par un apprenant A2, alors que d'autres formations lexicales moins transparentes ou moins fréquentes dans l'écriture de Zola peuvent représenter un obstacle supplémentaire.

Ces différents éléments ne sont cependant pas étudiés séparément. Ils sont toujours replacés dans leur contexte phrastique et discursif, car la difficulté d'un mot dépend

également de la manière dont il apparaît dans le texte et des relations qu'il entretient avec les autres éléments de la phrase.

Au-delà du choix des mots, les versions simplifiées révèlent également un ensemble d'opérations lexicales récurrentes. Certaines transformations prennent la forme d'un remplacement par un terme générique, par exemple lorsqu'un nom précis est substitué par une désignation plus large. D'autres consistent en une suppression d'adjectifs qualificatifs, en particulier lorsque ceux-ci participent à l'accumulation descriptive caractéristique du style naturaliste. On observe enfin des reformulations périphrastiques, où un mot jugé opaque est remplacé par une expression plus explicite. À ce niveau de l'analyse, nous nous limitons volontairement à la description de ces opérations, sans encore en tirer de conclusions sur leurs effets sémantiques ou stylistiques.

On aurait tort cependant de réduire la simplification au seul plan lexical. La syntaxe joue un rôle tout aussi déterminant. Nous avons donc examiné la longueur des phrases et leur éventuelle segmentation. Dans plusieurs cas, une phrase complexe du texte source est fragmentée en unités plus simples, ce qui modifie l'organisation interne de l'énoncé sans nécessairement en transformer le contenu informationnel. Ce phénomène apparaît de manière particulièrement nette lorsque de longues descriptions sont redistribuées sous forme de phrases brèves, par exemple lorsqu'un portrait dense issu d'*Au Bonheur des Dames* est reformulé en une succession d'énoncés courts du type : « Pépé, à douze ans, est déjà plus grand qu'elle. Il est toujours aussi gentil. Jean est très beau, avec ses cheveux blonds. » (Claustres, 2023) Ce découpage tend à alléger la charge cognitive immédiate, mais il transforme aussi la continuité phrastique qui soutient le rythme de l'écriture zolienne.

Cette transformation s'accompagne souvent d'un remaniement de la structure. Les constructions fondées sur la subordination tendent à être remplacées par des formes plus juxtaposées, notamment lorsque des propositions relatives ou circonstancielles deviennent des phrases autonomes. Les incises descriptives, fréquentes dans le texte source, peuvent être déplacées ou supprimées, ce qui influe sur la progression descriptive et sur la construction des images. Nous avons également relevé des

modifications plus discrètes, telles que la réduction des expansions du nom ou le réagencement de l'ordre des groupes syntaxiques, qui participent à une reconfiguration interne du discours.

Ces modifications ne relèvent pas uniquement d'un allègement grammatical. Elles touchent aussi à la dynamique même du texte, dans la mesure où la syntaxe constitue l'un des supports du rythme et de la densité descriptive propres à l'écriture zolienne. Modifier la structure des phrases revient donc à transformer la manière dont le monde représenté est donné à voir.

Toutefois, même si l'analyse linguistique est indispensable, elle ne saurait suffire à elle seule. Une phrase courte peut demeurer difficile lorsqu'elle concentre un lexique spécialisé ou des références implicites, tandis qu'une phrase longue peut rester lisible si sa structure est régulière et si ses enchaînements sont clairs. Il convient donc de nuancer l'idée selon laquelle la simplification formelle entraînerait mécaniquement une amélioration de la lisibilité.

Ces limites nous conduisent à dépasser une approche strictement descriptive pour interroger les procédés concrets par lesquels le texte est transformé. L'analyse linguistique prépare ainsi le terrain de l'étape suivante, consacrée à la typologie des procédés de simplification, qui permettra de mieux comprendre comment les modifications repérées sont produites et combinées.

2.3.2. Typologie des procédés de simplification

Après avoir noté certains changements au niveau du lexique et de la syntaxe (Axe 1), nous nous sommes interrogée un peu plus concrètement sur les processus par lesquels ces modifications surviennent. Il ne s'agit plus seulement de constater la différence entre les versions, mais de comprendre plus concrètement comment s'opère cette transformation. Parfois, on a l'impression que c'est subtil, presque invisible, mais en réalité, aucun écart n'est fortuit.

Dans cette perspective, nous avons retenu une typologie structurée autour de cinq procédés opératoires : la suppression, la substitution, la reformulation, l'explicitation et le découpage. Ce choix s'appuie sur les travaux consacrés à la reformulation et à la paraphrase (Martinot, 2010 ; Gala et al., 2020 ; Fuchs, 2020), ainsi que sur les

recherches plus anciennes portant sur la simplification textuelle et l'explicitation (Blum et Levenston, 1978 ; Blum-Kulka, 1986). Elle a été enfin réélaborée en fonction des contraintes propres à notre corpus.

Nous observons d'abord des cas de suppression, qui correspondent à l'effacement pur et simple de segments du texte source. Ces suppressions peuvent concerner des syntagmes isolés, des propositions subordonnées, voire des séquences descriptives plus étendues. Dans les versions simplifiées, nous remarquons souvent que certains détails culturels ou contextuels disparaissent, notamment lorsqu'ils ne sont pas jugés indispensables à la compréhension immédiate de la scène. Ce type d'intervention ne réduit pas seulement la longueur du texte : il change aussi la quantité d'informations et la richesse des références du discours. Il n'est pas exclu que l'appauvrissement de ces réseaux de références contribue à atténuer ce qui fait la singularité de l'écriture zolienne, en particulier sa capacité d'ancrer les personnages dans un environnement matériel et social fortement caractérisé. La suppression apparaît ainsi comme un procédé direct. Ses effets sur la représentation du monde narratif demeurent, en ce sens, loin d'être neutres.

À côté de ces effacements, les données font apparaître de nombreuses opérations de substitution. Celles-ci reposent sur le remplacement d'un élément linguistique par un autre qui est jugé plus accessible. Le phénomène est particulièrement visible au niveau lexical, lorsque des termes perçus comme spécialisés ou peu fréquents sont remplacés par des mots plus courants. C'est le cas, par exemple, lorsque le terme « arithmétique » du texte source est remplacé par le mot plus général « mathématique » dans la version éditoriale. Ce choix semble réduire l'effort de lecture pour le public ciblé, notamment les apprenants de niveau A2. La substitution peut également concerner la syntaxe, lorsqu'une structure complexe cède la place à une construction plus simple, sans que le contenu informationnel soit entièrement supprimé. Contrairement à la suppression, il ne s'agit donc pas de retrancher un segment, mais d'en proposer une forme alternative, censée réduire l'effort de traitement pour le lecteur.

Plus complexe à identifier, la reformulation se distingue des procédés précédents par son caractère plus global. Elle ne repose ni sur l'effacement pur ni sur un simple remplacement ponctuel, mais sur une reprise de l'énoncé sous une autre forme. Le contenu général reste souvent proche du texte source, tandis que l'organisation phrastique ou le mode d'expression se modifient plus largement. Dans certains cas, il devient difficile d'établir une correspondance précise entre les deux versions : on a alors davantage l'impression d'une réécriture partielle du passage. La frontière avec la substitution reste d'ailleurs parfois floue. Comme certaines transformations combinent plusieurs opérations à la fois, nous avons choisi d'accepter une part de chevauchement dans le codage plutôt que d'imposer des catégories trop rigides.

Nous retenons également l'explicitation comme procédé autonome, lorsque la version simplifiée rend explicites des éléments seulement suggérés dans le texte source. Cela peut passer par l'ajout d'un lien logique, la précision d'un référent ou le développement d'une relation implicite. Ce procédé tend généralement à réduire les zones d'ambiguïté ou d'inférence présentes dans le texte original, afin de faciliter l'accès au sens pour le lecteur.

Enfin, le découpage correspond à la segmentation d'une phrase longue ou dense en plusieurs phrases plus courtes. Les subordonnées sont fréquemment dissociées et les informations redistribuées de manière plus linéaire. Ce procédé vise surtout à alléger la charge syntaxique du texte. Mais cette fragmentation modifie aussi parfois son rythme. Certains mouvements de phrase ou effets de continuité deviennent moins perceptibles dans les versions simplifiées.

Afin de rendre ces phénomènes analysables, chaque modification observée entre le texte source et ses versions simplifiées est systématiquement repérée puis associée à l'un de ces procédés. Nous avons toutefois choisi de ne pas imposer une classification exclusive : lorsqu'une transformation combine, par exemple, une suppression partielle et une reformulation paraphrastique, elle peut être codée comme relevant de plusieurs catégories. Cette souplesse vise à respecter la complexité effective des opérations observées, plutôt qu'à forcer les données à entrer dans un schéma prédéfini.

Cette typologie nous semble utile pour deux raisons principales. Elle nous permet, en premier lieu, de comparer les pratiques de simplification mises en œuvre par un sujet humain et par un système d'intelligence artificielle. Il ne s'agit pas de dire qu'un texte est globalement plus simple ou non, mais de regarder concrètement quels types d'opérations sont vraiment utilisées. Parallèlement, elle offre un outil d'analyse qualitative des stratégies derrière les simplifications, en rendant visible les mécanismes de transformation privilégiés dans chaque version. En adoptant cette approche, nous n'en restons pas à une impression générale ou à un ressenti intuitif du résultat final ; nous essayons de construire une description raisonnée des mécanismes spécifiques par lesquels le texte se transforme.

2.3.3. Tension entre fidélité sémantique et littérarité

L'un des enjeux majeurs de notre protocole a été de tenir ensemble deux dimensions que le processus de simplification tend paradoxalement à dissocier : la fidélité sémantique et la littérarité. À ce stade de la réflexion, il nous a semblé difficile de vouloir tout observer à la fois ; c'est pourquoi nous avons arrêté nos critères d'observation sur de la sémantique et du style. Cette décision n'a pas été prise à la légère : il s'agissait d'un arbitrage méthodologique nécessaire, destiné à stabiliser les invariants narratifs — le maintien des relations logiques et la traçabilité des référents — tout en restant attentifs aux glissements de sens, parfois imperceptibles, qui surviennent lors d'une reformulation ou d'une explicitation excessive.

Nous avons retenu ces catégories car elles permettent d'observer plus précisément l'équilibre entre les éléments conservés et ceux qui sont transformés au cours de la simplification. Pour réaliser le codage, chaque extrait du corpus a été étudié de manière progressive. Nous avons d'abord repéré les éléments textuels qui étaient maintenus malgré la simplification, puis ceux dont la présence diminuait ou disparaissait, totalement ou partiellement. Cette démarche ne visait donc pas uniquement à relever les éléments supprimés. Elle cherchait surtout à comprendre comment les modifications apportées transforment l'organisation du texte, en identifiant les changements de sens, les déplacements ou les atténuations introduits par le processus de simplification. Ce travail de codage progressif nous a conduit à

élaborer une grille d'analyse opérationnelle. Les différents instruments d'analyse et de codage mobilisés dans cette recherche, dont cette grille constitue l'un des éléments centraux, sont présentés dans leur intégralité en *Annexe 4*. Chaque critère répondait à une question précise : qu'est-ce qui est conservé, qu'est-ce qui se transforme, et quels en sont les effets sur la réception ?

Cette volonté d'objectivation nous a également guidée dans le traitement de la littérarité, que nous n'avons pas considérée comme une impression subjective. Sans chercher à mesurer la qualité littéraire du texte, nous avons plutôt essayé d'identifier certains marqueurs récurrents de l'écriture zolienne, notamment les images, les accumulations descriptives ou encore certains effets rythmiques. Il ne s'agissait donc pas uniquement de noter des disparitions, mais aussi de suivre les déplacements ou les reconfigurations des effets stylistiques au fil de la simplification. Dans plusieurs cas, le contenu narratif reste compréhensible alors même que la texture stylistique du passage se transforme sensiblement.

Au cours du repérage des occurrences, nous avons notamment été attentive à certains phénomènes récurrents : la neutralisation d'adjectifs évocateurs, la réduction des accumulations descriptives ou encore l'explicitation de certains implicites. Ces transformations ne modifient pas toujours directement l'information narrative, mais elles changent parfois la manière dont le lecteur entre dans le texte ou perçoit l'univers représenté. Autrement dit, la conservation du sens ne garantit pas nécessairement le maintien de la même expérience de lecture. Cette observation nous a amenée à accorder une attention particulière à l'équilibre entre lisibilité et maintien d'une certaine densité stylistique.

Finalement, pour que ce protocole ne reste pas une simple intuition, nous avons structuré nos critères d'évaluation autour des deux indicateurs définis dans notre grille : d'une part, la sémantique et la réception (degré de conservation des invariants et glissements de sens) ; d'autre part, le style et la narration (typologie des pertes esthétiques et transformation de l'immersion perceptive). Le choix de ces critères répond à un objectif précis : permettre une analyse qui puisse être suivie et reproduite à partir des mêmes éléments. La première dimension fournit une mesure quantitative

de la conservation des informations narratives, tandis que la seconde permet d'observer plus finement les transformations liées au rythme du texte, aux images utilisées ou encore aux effets produits sur le lecteur. L'explicitation de ces critères permet ainsi de clarifier notre démarche d'analyse. Elle montre les différents aspects du texte pris en compte lors de l'étude de la simplification et permet d'évaluer non seulement les éléments conservés, mais aussi les modifications apportées à la dimension esthétique du texte source.

2.3.4. Évaluation de la lisibilité objective

Dans cette section, nous avons cherché à répondre à une question volontairement simple en apparence : que disent les chiffres ? Autrement dit, comment objectiver les effets de la simplification à l'aide d'indicateurs mesurables, sans pour autant réduire la lisibilité à une simple impression subjective. À ce stade de la réflexion méthodologique, il nous a semblé nécessaire de recourir à des outils standardisés, capables de fournir des résultats comparables et reproductibles, tout en restant adaptés aux spécificités linguistiques du français.

Afin d'analyser la lisibilité des textes, nous avons retenu deux indices complémentaires : l'indice de Flesch et celui de Kandel-Moles. Ces mesures sont complétées par une analyse du profil lexical réalisée avec l'outil Lextutor. L'association de ces trois approches répond à un double objectif. D'une part, elle permet de disposer d'indicateurs adaptés à l'étude d'un corpus en français ; d'autre part, elle facilite la mise en perspective de nos résultats avec les travaux antérieurs qui utilisent fréquemment l'indice de Flesch.

L'indice de Flesch (1948) constitue l'un des indicateurs de lisibilité les plus largement utilisés dans la littérature. Il repose sur l'idée que la difficulté d'un texte dépend en grande partie de deux facteurs formels : la longueur moyenne des phrases et la longueur moyenne des mots. Plus les phrases sont longues et plus les mots comportent de syllabes, plus le traitement cognitif est supposé être coûteux pour le lecteur. Le score obtenu se situe généralement sur une échelle allant de 0 à 100, un score élevé correspondant à une lisibilité plus facile.

Toutefois, l'indice de Flesch a été conçu initialement pour l'anglais dont la structure morphologique et syllabique diffère sensiblement du français. C'est donc l'adaptation proposée par Kandel and Moles (1958) pour le français que nous avons privilégié comme mesure principale de la lisibilité de notre corpus, afin de respecter la réalité morphologique du français, tout en conservant l'indice de Flesch comme point de référence pour assurer une comparaison internationale. La formule Kandel-Moles utilisée est la suivante :

$$\text{FRE} = 207 - (1,015 \times \text{ASL}) - (73,6 \times \text{ASW})$$

Où :

- FRE (Flesch Reading Ease) indique l'indice de lisibilité du texte ;
- ASL (Average Sentence Length) correspond à la longueur moyenne des phrases, exprimée en nombre de mots par phrase ;
- ASW (Average Syllables per Word) correspond au nombre moyen de syllabes par mot.

Cette formule accorde un poids important à la variable syllabique, ce qui est particulièrement pertinent pour le français, langue dans laquelle la longueur des mots en syllabes est statistiquement plus élevée que dans l'anglais. Nous avons donc retenu l'indice de Kandel-Moles comme mesure principale de la lisibilité de notre corpus, tout en maintenant l'indice de Flesch comme point de référence. Le recours conjoint à ces deux indices nous permet de croiser une approche adaptée aux spécificités du français avec une échelle largement utilisée dans la littérature existante.

Le traitement de chaque extrait (VO, VE, VIA) a soulevé plusieurs difficultés d'ordre méthodologique, en particulier en ce qui concerne le calcul du nombre de syllabes selon la méthode de Kandel-Moles. Nous avons d'ailleurs dû trancher sur le cas des e-muets, souvent problématiques dans le décompte syllabique des versions littéraires. Les liaisons ont aussi posé question : faut-il compter le « s » final prononcé ou muet selon le contexte ? Pour chacun des textes, nous avons procédé à un comptage systématique du nombre total de phrases, de mots et de syllabes, étape indispensable à l'obtention des valeurs de ASL et de ASW. Ces valeurs ont ensuite été intégrées à la fois dans la formule de Kandel-Moles ainsi que dans celle de Flesch afin de

produire un score de lisibilité pour chaque version du même passage. L'application de ce protocole identique à l'ensemble des extraits vise à assurer la traçabilité et la reproductibilité des mesures obtenues.

Afin de rendre ces scores interprétables, nous nous appuyons sur une grille de lecture des indices de lisibilité, qui associe des plages de scores à des niveaux de difficulté. Les scores élevés correspondent à des textes très faciles à lire, tandis que les scores faibles indiquent une forte difficulté de lecture. Cette grille nous permet de traduire les valeurs numériques en catégories interprétatives telles que « facile », « moyen », « difficile », et de situer les différentes versions d'un même extrait les uns par rapport aux autres. Il ne s'agit donc pas seulement de produire des chiffres, mais de disposer d'un outil permettant de comparer objectivement les effets de la simplification. Le tableau ci-dessous propose une correspondance entre les valeurs de l'indice de Flesch et des niveaux de lisibilité généralement admis dans la littérature. Il s'agit ici de la version destinée à l'anglais, et nous l'utilisons comme point de référence indirect, car les scores de Kandel-Moles pour le français peuvent présenter de légères différences par rapport à l'échelle de 0 à 100 prévue pour l'anglais.

Tableau 3. *Interprétation des scores de Flesch adaptée d'après Flesch (1948)*

Score Flesch	Difficulté	Exemple
90 – 100	Extrêmement facile	Textes de niveau élémentaire
80 – 90	Très facile	Bandes dessinées
70 – 80	Facile	Annonces promotionnelles
60 – 70	Moyen	Médias grand public
50 – 60	Assez difficile	Presses de référence
30 – 50	Difficile	Ouvrages académiques
0 – 30	Très difficile	Publications de recherche spécialisée

Cependant, la lisibilité ne dépend pas uniquement de la structure syntaxique ni de la longueur des mots. Elle est également liée à la nature du lexique employé. C'est la

raison pour laquelle nous avons complété l'analyse par un recours à *The Compleat Lexical Tutor*, souvent appelé *Lextutor*, un outil en ligne qui permet d'établir le profil lexical des textes en fonction de la fréquence des unités lexicales.

Lextutor répartit les mots d'un texte en plusieurs bandes de fréquence, généralement désignées par les étiquettes K1, K2, K3, ... et Off-list. Chaque catégorie correspond à un degré différent de fréquence et de familiarité pour le lecteur :

- K1 correspond aux mots les plus fréquents de la langue, issus du premier millier lexical ; il s'agit principalement de vocabulaire courant et fonctionnel.
- K2 groupe des mots également fréquents, mais appartenant à un second niveau de fréquence, souvent plus variés sur le plan sémantique.
- K3 concerne des mots moins fréquents, susceptibles de poser davantage de difficultés aux lecteurs peu avancés.
- Au-delà de K3, il existe également d'autres catégories (K4, K5, etc) qui regroupent des mots de fréquence encore plus faible ; ces catégories reflètent un lexique rare ou spécialisé.
- Off-list désigne les unités absentes des listes de fréquence de base ; cette catégorie est généralement associée à un lexique rare, spécialisé ou littéraire.

Ce découpage permet d'évaluer la proportion relative de mots fréquents et de mots rares dans chaque texte, et ainsi d'approcher la complexité lexicale de manière quantitative.

Dans le cadre de notre étude, nous retenons principalement :

- la catégorie K1 comme indicateur central du lexique fréquent, supposé le plus accessible pour des apprenants de niveau A2 ;
- la catégorie K2 comme zone d'extension lexicale, correspondant à des unités moins fréquentes mais susceptibles d'être partiellement connues par des apprenants en fin de niveau A2 ou en transition vers le niveau B1 ;
- les catégories K3 et Off-list comme indicateurs d'un lexique peu fréquent ou potentiellement difficile.

Le choix de ces indicateurs s'appuie sur une comparaison entre les bandes de fréquence fournies par Lextutor et les estimations du répertoire lexical correspondant au niveau A2 du CECRL. Selon ces estimations, un apprenant de ce niveau disposerait d'un vocabulaire réceptif d'environ 1000 à 1200 mots et d'un vocabulaire productif situé autour de 600 à 800 mots. Dans cette logique, la catégorie K1 occupe une place centrale puisqu'elle rassemble les 1000 mots les plus fréquents de la langue française. Elle correspond ainsi au socle lexical le plus susceptible d'être connu ou reconnu par un apprenant A2. Pour notre analyse, nous considérons qu'un texte présente une accessibilité lexicale relativement satisfaisante lorsque la part des mots appartenant à K1 atteint environ 85% ou davantage. Ce pourcentage a été retenu comme un repère méthodologique : en dessous de ce niveau, l'augmentation des mots moins fréquents peut entraîner une charge lexicale plus importante et rendre l'accès au sens plus difficile pour les apprenants. Il convient toutefois de préciser que ce seuil ne correspond pas à une norme officielle. Il sert uniquement de point de comparaison entre les différentes versions étudiées.

La catégorie K2 correspond au deuxième millier de fréquence. Elle représente un vocabulaire moins courant que celui de K1, mais qui peut encore apparaître dans des textes destinés à un niveau A2, notamment comme une ouverture vers un lexique plus large. Dans notre étude, une proportion située entre 7% et 10% est retenue comme indication permettant d'interpréter la présence de ce type de vocabulaire.

En revanche, les catégories K3 et Off-list signalent un lexique plus rare, spécialisé, littéraire ou absent de la base lexicale de référence. Leur présence est interprétée comme un indice potentiel de complexité lexicale accrue ; une proportion inférieure à 5% est retenue à titre indicatif comme repère d'analyse.

Cette répartition nous permet de comparer les versions d'un même extraits (VO, VE, VIA) en observant les éventuels déplacements du lexique rare vers un lexique plus fréquent, ce qui constitue un indice de simplification lexicale.

Ainsi, *Lextutor* ne sert pas uniquement à décrire le lexique, mais à compléter les indices de lisibilité de type de Flesch ou Kandel-Moles en introduisant une dimension lexicale que ces formules ne prennent pas directement en compte. Un texte peut, en

effet, présenter des phrases courtes tout en mobilisant un vocabulaire peu fréquent, ce qui peut en affecter la compréhension réelle. Nous estimons donc que l'analyse combinée de la structure syntaxique et de la distribution lexicale permet d'obtenir une évaluation plus fine de la lisibilité effective des textes.

Les variables mesurées dans cette section sont donc de deux ordres. D'une part, des variables structurelles : longueur moyenne des phrases (ASL) et longueur moyenne des mots en syllabes (ASW). D'autre part, des variables lexicales : distributions des mots selon les bandes de fréquence K1, K2, K3, et Off-list, ainsi que la proportion globale de mots fréquents.

Il nous a cependant semblé nécessaire de ne pas considérer ces indices comme des verdicts définitifs. Même s'ils mettent en lumière certaines caractéristiques formelles des textes, ils ne disent rien de ce que le lecteur ressent réellement en lisant, ni de la manière dont le style ou les effets discursifs influencent son expérience. C'est un constat qui nous a paru évident en manipulant les extraits, en comparant VO, VE et VIA, et en observant les petites nuances que les chiffres ne capturent pas.

Dans cette recherche, nous avons donc choisi de confronter systématiquement ces résultats quantitatifs aux analyses linguistiques menées dans les sections précédentes, puis de prolonger ce travail par une interprétation plus qualitative et humaine dans la section suivante. Ce regard permet de saisir concrètement comment les transformations affectent la lisibilité et la compréhension : les indices ne sont qu'un point de départ, un outil pour orienter l'observation, mais c'est l'analyse attentive qui révèle les subtilités de lecture, les glissements de sens et les ajustements stylistiques qui échappent aux chiffres.

En somme, il ne s'agit pas seulement de mesurer et de constater, mais de comprendre et de documenter ce que la simplification fait à l'expérience de lecteur — ce mélange fragile entre lisibilité et richesse stylistique, que chaque texte rend unique.

2.4. Étude empirique : le regard de l'expert FLE

2.4.1. Objectifs de l'évaluation par perception humaine

Dans les sections précédentes, nous avons présenté le processus d'évaluer la lisibilité des textes simplifiés à partir de plusieurs angles d'analyse. Certains résultats

proviennent d'indices de lisibilité calculés automatiquement ; d'autres reposent déjà sur une observation qualitative menée par nous-même, portant notamment sur le lexique, la structure syntaxique ou encore les procédés de simplification employés. Cette double approche nous permet d'obtenir une première compréhension du phénomène étudié. Pourtant, une question nous est alors venue naturellement : comment ces textes se comportent-ils réellement face au regard d'experts du FLE, qui sont habitués à guider des apprenants ?

Les indices de lisibilité offrent bien sûr un premier repère pour estimer la complexité linguistique d'un texte. Ils reposent sur des critères relativement objectifs — longueur moyenne des phrases, nombre de syllabes par mot, densité lexicale — et permettent d'obtenir rapidement une estimation globale de la difficulté d'un passage. Par exemple, nous avons observé qu'un texte avec des phrases courtes et des mots fréquents produisait systématiquement un score élevé. Au contraire, lorsqu'un passage accumule subordonnées et termes rares, le score chute, signalant un texte plus difficile.

Ces indicateurs reposent cependant sur une logique essentiellement formelle. Nous remarquons donc qu'ils ne prennent pas toujours en compte certaines dimensions plus fines du discours. La cohérence interne d'un texte, par exemple, échappe largement à ce type de calcul. Deux phrases peuvent être très courtes, parfaitement simples d'un point de vue grammatical ; si la transition entre elles reste floue, la compréhension globale devient soudain moins évidente. De la même manière, la fluidité de la lecture ne dépend pas uniquement de la longueur des phrases. Une succession de phrases très brèves peut produire un effet saccadé. On pense par exemple à une reformulation du type : « Il entre dans le magasin, Il regarde les vitrines. Il voit beaucoup de choses. Il est surpris. » Les phrases sont simples, mais la lecture sonne mécanique, presque artificielle. Ce type de dissonance ne peut pas être détecté un indice chiffré.

Nous avons aussi remarqué que la clarté de l'organisation des idées constitue une autre dimension difficile à mesurer automatiquement. Un texte peut présenter un score de lisibilité favorable, mais laisser le lecteur incertain quant au fil logique. Dans certains passages générés par l'IA, les phrases restent grammaticalement correctes,

parfois même très simples, et pourtant la progression des informations semble décalée, comme si le texte perdait une partie de sa cohérence narrative. Après avoir observé ce phénomène, nous avons compris que seule une lecture attentive par un expert pouvait révéler ces subtilités.

C'est là qu'intervient notre analyse qualitative initiale. Nous avons déjà tenté d'aller plus loin grâce à une observation qualitative du corpus : choix lexicaux, simplifications syntaxiques, transformations sémantiques ou encore maintien — ou perte — de certaines nuances du texte d'origine. Cette première lecture analytique permet de repérer plusieurs phénomènes intéressants, notamment les tensions possibles entre simplification linguistique et préservation du sens littéraire. Mais malgré ces observations, nous avons encore besoin d'un point de vue extérieur, concret, enraciné dans l'expérience pédagogique. C'est pour cela que nous avons sollicité des enseignants du FLE.

Dans le contexte de la didactique du FLE, ces experts apportent un regard unique. Leur expérience en classe les amène à détecter rapidement ce qui est compréhensible pour un niveau donné et ce qui peut déstabiliser les apprenants. Cette compétence ne repose pas uniquement sur des critères formels ; elle s'appuie aussi sur une expérience concrète de la lecture en classe, de la réaction des apprenants, des difficultés de compréhension qui apparaissent parfois là où on ne les attend pas. Il suffit de penser à certaines expressions culturellement marquées — une allusion à un grand magasin parisien, une remarque ironique sur les habitudes mondaines, ou encore une formule idiomatique dans le texte, et c'est précisément ce qui rend leur jugement si précieux. La machine peut simplifier la syntaxe ou remplacer un mot rare, mais elle ne saisit pas toujours les nuances culturelles ni les implicites qui structurent le sens.

Le regard des experts devient alors, d'après nous, particulièrement pertinent. Leur évaluation permet d'apprécier la lisibilité d'un texte selon plusieurs dimensions complémentaires : non seulement la facilité de compréhension globale, mais aussi la fluidité de la lecture, la cohérence du discours ou encore l'adéquation au niveau linguistique visé. La naturalité de l'expression joue également un rôle crucial. Un texte simplifié peut être grammaticalement correct tout en donnant une impression

d'artificialité ; ce type de perception, difficile à quantifier automatiquement, apparaît souvent très rapidement aux yeux d'un enseignant expérimenté. Leur perception apporte donc une dimension pédagogique que nous ne pouvions atteindre par des mesures chiffrées seules.

Loin de vouloir disqualifier les mesures automatiques, nous pensons que l'évaluation humaine affine les contours de l'analyse. Les indices de lisibilité fournissent une première indication chiffrée, notre lecture qualitative menée plus haut identifie certaines stratégies de simplification, et l'expertise FLE permet enfin de juger l'impact réel sur l'apprenant.

Dans ce cadre, notre étude perceptive poursuit plusieurs objectifs précis. Nous souhaitons recueillir l'appréciation d'enseignants ou de spécialistes de FLE sur des extraits simplifiés par l'humain et par l'IA. Les participants sont invités à évaluer ces passages selon plusieurs critères liés à la lisibilité et à la qualité discursive. Nous voulons comprendre si, selon eux, les textes IA sont lisibles et fluides, s'ils paraissent naturels ou non, et quelles différences émergent dans la cohérence du discours. Nous cherchons surtout à voir quelles dimensions attirent particulièrement l'attention des enseignants et comment elles se manifestent dans la lecture.

En croisant ces observations avec les analyses quantitatives et qualitatives réalisées précédemment, nous espérons ainsi obtenir une vision plus nuancée du phénomène étudié. La lisibilité d'un texte ne se réduit jamais à un chiffre ; elle se construit dans l'interaction entre une structure linguistique, une organisation discursive et la perception d'un lecteur. C'est cette rencontre, intime et concrète, que notre étude cherche à éclairer.

2.4.2. Échantillonnage : profil des participants

L'évaluation des textes simplifiés a été confiée à des enseignants de français langue étrangère, dans la mesure où cette recherche vise moins à mesurer la seule facilité de lecture qu'à apprécier l'intérêt de ces textes dans une perspective pédagogique. Habités à sélectionner et à exploiter des supports destinés à différents profils d'apprenants, les enseignants de FLE sont en mesure d'en évaluer à la fois la lisibilité, l'adéquation au niveau visé et les possibilités d'exploitation en classe.

Les participants avaient tous un lien avec le FLE, soit par leur formation, soit par leur expérience professionnelle. Le questionnaire demandait également depuis combien de temps ils enseignaient le FLE. Cette donnée avait pour seul objectif de décrire le profil de l'échantillon et de mettre les réponses en perspective selon leur expérience. Le recrutement s'est déroulé sur plusieurs semaines. Le questionnaire a d'abord été diffusé dans différents groupes Facebook réunissant des enseignants de FLE. Plusieurs enseignants-chercheurs de l'Université de Hanoï et de l'Université catholique de Louvain ont ensuite accepté de le transmettre à leurs collègues ainsi qu'à leurs réseaux professionnels. Ce mode de diffusion nous a permis de réunir un échantillon composé d'enseignants exerçant dans des contextes variés tout en partageant une expertise commune en didactique du FLE.

Au total, vingt-huit questionnaires complets ont été retenus pour l'analyse. Cet effectif ne permet pas de généraliser les résultats à l'ensemble des enseignants de FLE, mais telle n'était pas l'ambition de cette étude. Celle-ci s'inscrit dans une démarche exploratoire cherchant avant tout à recueillir des appréciations argumentées de professionnels susceptibles d'apporter un éclairage sur les qualités pédagogiques des textes proposés.

2.4.3. Enquête à l'aveugle et codifications des extraits

Dans cette étude, nous avons conçu l'évaluation perceptive de façon à limiter au maximum les biais d'interprétation. Il nous semble important que les participants jugent les textes uniquement sur leurs qualités intrinsèques et non sur leur origine supposée. Pour cette raison, nous avons opté pour l'enquête en aveugle, méthode fréquemment utilisée dans les recherches expérimentales pour neutraliser les attentes des évaluateurs. Au moment de répondre au questionnaire, les participants ne disposent d'aucune information sur l'origine des textes. Ils ignorent si chaque extrait provient d'une adaptation éditoriale ou d'une génération par IA. Ils lisent ensuite les extraits un à un avant de les évaluer selon différents critères, comme la clarté, la fluidité, la naturalité ou l'adaptation au niveau des apprenants. L'objectif est qu'ils fondent leur appréciation sur le texte lui-même, sans être influencés par son origine.

La question centrale n'est plus « qui a écrit ce texte ? », mais : est-il agréable et compréhensible à lire ? Est-il adapté à une classe de FLE de niveau A2 ?

Ce choix de méthodologie nous paraît particulièrement important dans un contexte où les représentations autour de l'IA restent souvent très marquées. Les enseignants peuvent avoir des attentes ou des préjugés : certains sont fascinés par les outils numériques, d'autres restent méfiants ou sceptiques. Si l'origine des textes était connue, les enseignants pourraient être influencés au moment de l'évaluation. La seule mention d'une version générée par l'IA pourrait conduire certains à porter un jugement plus sévère, tandis que d'autres pourraient, au contraire, se montrer plus indulgents par intérêt pour cette technologie. L'anonymisation des extraits avait précisément pour objectif d'éviter ce type de biais et de recueillir des évaluations fondées sur les caractéristiques des textes eux-mêmes.

Les extraits ont donc été codifiés avant d'être intégrés au questionnaire. Chaque texte est identifié par un code neutre — par exemple *Extrait 1*, *Extrait 2*, ... — afin d'éviter toute identification indirecte. Les participants ne voient pas non plus les extraits originaux de l'œuvre, et chaque texte est évalué comme un document autonome.

Les extraits sélectionnés pour le questionnaire (voir *Annexe 3*) proviennent du corpus analysé dans les sections précédentes. Toutefois, nous n'avons pas retenu l'ensemble des passages étudiés, car une telle option aurait rendu le questionnaire inutilement long. Lire un grand nombre d'extraits demande du temps et aussi de la concentration ; lorsque les textes deviennent trop nombreux, la fatigue peut s'installer et la qualité des réponses diminue. Nous avons donc préféré procéder à une sélection plus restreinte. En limitant le nombre de textes, nous cherchons au contraire à maintenir une attention réelle sur chaque extrait. Le participant peut prendre le temps de lire, de relire éventuellement, puis de formuler un jugement réfléchi. Ce dispositif permet également d'observer la lisibilité perçue des textes dans une situation de lecture relativement autonome, comparable à celle que pourraient rencontrer les apprenants lorsqu'ils découvrent un texte simplifié en classe.

Sélectionner les passages représentatifs n'a pas été simple. Nous avons tenté de choisir les extraits les plus pertinents, essayant de représenter différents types de

simplification, tout en tenant compte de la longueur des textes. Certains passages édités par des humains peuvent ne contenir qu'une seule phrase, tandis que la version IA du même passage peut compter plusieurs phrases pour expliciter ou reformuler les idées. Il nous semble donc compliqué de trouver un équilibre : comment éviter de privilégier systématiquement les textes courts ou longs et s'assurer que chaque extrait permette une évaluation juste et réaliste. Cette sélection reste en partie subjective, mais reflète notre souci de représentativité et de pertinence pédagogique.

La question de l'ordre de présentation mérite également d'être précisée. Ceux-ci ne sont pas regroupés par type de simplification. Les participants ne lisent donc pas d'abord toutes les versions humaines, puis toutes les versions générées par l'IA. Les textes apparaissent dans un ordre aléatoire. Ce choix vise à éviter un phénomène fréquent dans les questionnaires comparatifs : la tendance à comparer directement deux textes qui se suivent. Si deux versions du même passage apparaissaient l'une après l'autre, le lecteur serait immédiatement tenté de les confronter. En dispersant les extraits dans le questionnaire, nous limitons ce type de comparaison immédiate. Cette organisation permet aussi de limiter les effets liés à la fatigue ou à la baisse d'attention. Lorsqu'un enseignant lit plusieurs extraits à la suite, il est possible que son niveau de concentration varie au cours du questionnaire. Les premiers textes sont parfois lus plus attentivement que les derniers. En faisant varier l'ordre de présentation des extraits, ce risque est réparti entre les différentes versions et influence moins les résultats.

Le questionnaire a été administré au moyen d'un formulaire *Google Forms*. Ce choix offrait une plus grande souplesse aux participants, qui pouvaient répondre au moment qui leur convenait, sur ordinateur ou sur téléphone. Les réponses étaient enregistrées automatiquement, ce qui facilitait leur traitement par la suite. Chaque participant consultait les extraits proposés avant de les évaluer à partir des différents critères de lisibilité et de compréhension retenus pour l'étude.

2.4.4. Élaboration du questionnaire et des échelles d'appréciation

Notre questionnaire ne se réduit pas à une simple succession de questions. Sa construction a demandé plusieurs ajustements, notamment pour trouver un équilibre

entre précision méthodologique et temps de réponse raisonnable pour les participants. Puisque les enseignants sollicités exercent souvent une activité professionnelle chargée, il paraissait difficile de leur proposer un dispositif trop long. Nous avons ainsi cherché à limiter la durée de passation tout en conservant suffisamment d'éléments pour permettre une évaluation pertinentes des extraits. Dans sa forme finale, le questionnaire peut être complété en environ quinze minutes, durée qui inclut la lecture des passages proposés ainsi que l'évaluation des différents critères. Nous avons également tenu compte du risque de fatigue : en limitant la durée de l'enquête, nous voudrions que chaque participant conserve une attention soutenue sur tous les extraits et fournisse des évaluations réfléchies, sans que la lassitude n'influence ses réponses.

Avant d'entrer dans l'évaluation proprement dite, nous avons choisi de recueillir quelques informations contextuelles. Les participants indiquent s'ils sont locuteurs natifs ou non natifs du français et précisent leur éventuelle expérience dans l'enseignement du FLE. Même si ces éléments ne constituent pas l'objet principal de l'étude, ils permettent de mieux interpréter les réponses obtenues et d'observer, si possible, si certains profils tendent à percevoir différemment la lisibilité des textes.

Après cette courte étape introductive, le questionnaire conduit les participants vers la partie centrale : la lecture et l'évaluation des extraits simplifiés. Chaque enseignant découvre une série de courts passages et doit porter une appréciation sur plusieurs aspects liés à la lisibilité et à l'usage pédagogique possible des textes.

Le choix des critères n'a pas été immédiat. Plusieurs dimensions auraient pu être retenues, et il a fallu en réduire le nombre afin d'éviter un questionnaire trop long. Nous avons finalement conservé quelques aspects jugés essentiels et de nouvelles dimensions centrées sur l'enseignement. Ces différents critères ont été traduits en items d'évaluation dans le questionnaire, présentés de manière synthétique dans le tableau ci-dessous.

Tableau 4. *Dimension expérimentale : critères d'évaluation du questionnaire*

Critère	Formulation dans le questionnaire
Pertinence d'intégration dans les supports pédagogiques	Pourriez-vous retenir ce texte et l'intégrer dans votre corpus pour le travailler en classe ?
Clarté pour les apprenants de niveau A2	Le texte est-il compréhensible pour ce niveau ?
Potentiel pédagogique	Peut-il servir de support pour enseigner un point linguistique, littéraire ou culturel ?
Capacité à susciter l'interaction	Peut-il susciter des échanges entre apprenants ?

Nous avons choisi de réduire l'attention portée à des dimensions purement liées à la qualité littéraire ou stylistique afin que de concentrer l'évaluation sur l'utilité pédagogique et la lisibilité pour des apprenants de niveau A2. Ces critères permettent d'observer différentes facettes de l'expérience de lecture et de l'usage possible en classe.

Pour chaque dimension, les participants attribuent une note à l'aide d'une échelle de Likert (1932) en cinq points, allant de 1 à 5.

Tableau 5. *Échelle de Likert pour l'évaluation des extraits simplifiés*

Note	1	2	3	4	5
Signification	Très faible	Faible	Moyen	Bon	Très bon

Ce choix de cette échelle a fait l'objet de quelques hésitations lors de la conception du questionnaire. Nous n'avons pas retenu une échelle à quatre points, car elle aurait conduit les participants à choisir systématiquement un avis positif ou négatif. Or, dans notre étude, certains jugements pouvaient être plus nuancés, notamment lorsqu'il s'agissait d'évaluer la lisibilité ou le potentiel pédagogique des textes. L'échelle à cinq points offrait donc une possibilité de réponse intermédiaire (« moyen »), qui

nous a semblé plus adaptée. Elle permet également d'attribuer une valeur numérique à chaque réponse tout en laissant une certaine souplesse dans l'évaluation.

Les notes chiffrées ne suffisent toutefois pas toujours à rendre compte de l'expérience réelle de lecture et de l'aptitude des textes à être exploités en classe. Nous avons donc proposé une question ouverte à la fin du questionnaire. Les participants peuvent ajouter quelques remarques personnelles sur les textes qu'ils viennent de lire, s'ils le souhaitent. Ces observations concrètes permettent d'éclairer les résultats numériques et de comprendre la perception des enseignants quant au potentiel pédagogique de chaque extrait.

Le questionnaire se termine par une question consacrée à l'usage de l'IA dans l'enseignement. Les participants sont invités à imaginer qu'un des textes évalués a été généré par une IA, puis à indiquer s'ils seraient prêts à l'utiliser avec leurs apprenants. L'objectif est d'explorer leur perception de la valeur pédagogique de ces textes ainsi que leur degré d'acceptation de l'IA comme outil d'aide à l'enseignement. Le questionnaire a été diffusé sur Google Forms. Les réponses ont été recueillies directement sur la plateforme. Aucune donnée permettant d'identifier les participants n'a été collectée. Seules des informations générales, comme le statut de locuteur natif ou non et l'expérience en FLE, ont été demandées. Les adresses e-mail et les autres informations d'identification n'ont pas été enregistrées, de manière à garantir l'anonymat des participants.

Une fois le questionnaire terminé, les réponses sont exportées depuis Google Forms, et nous les avons organisées dans un tableau pour pouvoir les analyser plus facilement. La participation était entièrement volontaire : chacun pouvait décider de répondre ou non, et arrêter le questionnaire à tout moment sans aucune conséquence. Cette précaution, simple mais essentielle, vise à ce que les enseignants se sentent libres de donner un avis sincère et réfléchi.

2.5. Traitement des données

Les données recueillies ont fait l'objet de traitements distincts, chaque catégorie répondant à un objectif d'analyse spécifique. Ce n'est qu'après cette première étape qu'elles ont été rapprochées dans une analyse d'ensemble.

Pour les données quantitatives, le travail s'est appuyé sur les différentes mesures de lisibilité calculées pour chacun des extraits du corpus. Les indices de Flesch, de Kandel-Moles, la longueur moyenne des phrases (ASL), la longueur moyenne des mots (ASW), ainsi que les résultats obtenus avec Lextutor, ont été rassemblés dans plusieurs tableaux comparatifs. Cette présentation permettait de visualiser plus facilement les écarts entre le texte original, la version éditoriale et la version générée par ChatGPT.

Le traitement des données qualitatives a suivi une autre démarche. Chaque extrait a d'abord été relu en parallèle avec les deux autres versions afin de repérer les transformations apportées au texte. Les procédés observés ont ensuite été répartis selon la grille d'analyse retenue dans cette étude : suppression, substitution, reformulation, explicitation et découpage des phrases. Au fil de cette lecture, une attention particulière a également été accordée au maintien du sens ainsi qu'aux éléments stylistiques qui restaient présents ou qui disparaissaient au cours de la simplification.

Les réponses recueillies grâce au questionnaire (voir *Annexe 8*) ont été examinées séparément. Les réponses aux questions fermées ont d'abord été récupérées depuis *Google Forms*, puis traitées dans un tableur pour calculer la moyenne de chaque critère. Les résultats sont présentés sous forme de tableaux et de graphiques. Pour les questions ouvertes, nous avons réalisé une analyse qualitative de contenu en suivant les recommandations de Paillé and Mucchielli (2021). Les commentaires ont été relus à plusieurs reprises avant d'être regroupés par thèmes. Cette démarche nous a permis d'identifier les principaux éléments relevés par les enseignants, qu'ils concernent les points forts des textes, leurs limites ou leurs possibilités d'exploitation pédagogique. La dernière étape consistait à mettre ces différents résultats en perspective. Les observations issues de l'analyse linguistique ont été comparées aux indices de lisibilité, puis aux évaluations recueillies auprès des enseignants. L'objectif n'était pas de considérer qu'une source de données était plus fiable qu'une autre. Il s'agissait plutôt d'observer dans quelle mesure elles conduisaient aux mêmes constats ou, au contraire, faisaient apparaître des différences méritant d'être examinées plus

attentivement. C'est sur ce croisement des analyses quantitative, qualitative et expérimentale que repose la démarche de triangulation adoptée dans cette recherche.

CHAPITRE 3 : ANALYSE DES RÉSULTATS ET DISCUSSION

3.1. Analyse quantitative de la lisibilité objective

Cette section regroupe les résultats des mesures quantitatives réalisées sur les trois versions du corpus. Les indicateurs analysés permettent d'évaluer l'impact des procédés de simplification sur la structure syntaxique et les caractéristiques lexicales des textes. Nous examinerons d'abord les indicateurs structurels, avant d'analyser plus en détail les différentes mesures statistiques retenues.

3.1.1. Comparaison des indicateurs structurels (ASL et ASW)

Le tableau ci-dessous permet de comparer globalement les transformations observées dans le corpus à partir de deux indicateurs principaux : la longueur moyenne des phrases (ASL – *Average Sentence Length*) et le nombre moyen de syllabes par mot (ASW – *Average Syllables per Word*).

Tableau 6. *ASL et ASW moyens des 14 extraits du corpus*

Version	ASL	ASW
VO	23	1,55
VE	9,3	1,43
VIA	10,3	1,51

Les écarts observés pour l'ASL sont remarquables : alors que la moyenne atteint 23 mots par phrase dans la version originale, cette valeur diminue nettement dans les deux versions simplifiées (9,3 pour la version éditoriale et 10,3 pour la version générée par ChatGPT). Cette différence nous suggère que la simplification repose d'abord sur une fragmentation importante de la structure phrastique.

Les moyennes globales se retrouvent également à l'échelle des extraits. Les écarts deviennent particulièrement visibles lorsque l'on compare les situations les plus extrêmes du corpus. Si les VO de l'extrait 2 et 6 dépassent les 39 mots par phrase en moyenne (39,8 et 41,3), certaines adaptations éditoriales descendent sous le seuil des 6 mots, comme dans les extraits 4 et 12 (5,6 et 4,2) (voir *Annexe 5*).

Cette réduction est visible dans l'ensemble du corpus, où les quatorze extraits présentent systématiquement un ASL inférieur à celui du texte source, quelle que soit la version étudiée. Les écarts ne sont pas toujours identiques, ce qui indique une application non uniforme des stratégies de simplification, mais la tendance reste constante.

Nous pouvons aussi observer que les résultats de la VIA suivent le même mouvement de réduction, mais avec une régularité légèrement moindre. Les valeurs demeurent systématiquement inférieures à celles de la VO tout en restant, dans plusieurs cas, légèrement supérieures à celles observées dans la VE. Cette position intermédiaire reflète une simplification essentiellement syntaxique, mais moins radicale que celle de la VE.

L'évolution de l'ASW nous présente toutefois une dynamique distincte. Les variations existent, mais elles sont moins marquées que celles observées pour l'ASL, ce qui montre que la simplification lexicale ne passe pas principalement par une modification de la longueur ou de la complexité phonologique des mots. La moyenne passe ainsi de 1,55 dans la VO à 1,43 dans la VE. La VIA affiche une moyenne de 1,51, proche de celle du texte source.

Quelques exemples permettent d'illustrer cette différence. Contrairement à la VE qui enregistre des baisses notables (jusqu'à 1,26 pour l'extrait 8), la VIA maintient des valeurs très similaires au texte source : l'ASW passe par exemple de 1,59 à 1,60 dans l'extrait 2 et de 1,56 à 1,57 dans l'extrait 4.

Ces premiers résultats montrent surtout un contraste marqué entre les versions en ce qui concerne la structure des phrases, tandis que l'ASW varie de manière plus limitée.

3.1.2. Comparaison des scores de lisibilité (Flesch et Kandel-Moles)

On ne peut nier que les indicateurs ASL et ASW permettent d'observer certaines caractéristiques du texte, mais ils ne donnent qu'une vision partielle de sa difficulté. Les scores de Flesch et de Kandel-Moles nous proposent une mesure plus globale en prenant en compte la longueur des phrases et celle des mots. Cette moyenne globale est détaillée dans le tableau ci-dessous : les scores des indices de Flesch et de Kandel-

Moles y sont identiques pour l'ensemble du corpus, ce qui justifie leur interprétation conjointe.

Tableau 7. *Scores de lisibilité des versions du corpus*

Version	Score de lisibilité moyen	Difficulté
VO	69,5	Moyen
VE	92,2	Extrêmement facile
VIA	85	Très facile

Pour l'ensemble du corpus, la version originale obtient une moyenne de 69,5. Sur les échelles de Flesch et Kandel-Moles, cette valeur correspond à un niveau de lisibilité moyen pour les natifs, ce qui représente un obstacle réel pour des apprenants de niveau A2. Les deux versions simplifiées présentent des scores nettement plus élevés : la VE atteint une moyenne de 92,2 et la VIA une moyenne de 85. Cet écart suggère que la simplification éditoriale réduit davantage les difficultés de lecture que la simplification produite par l'IA. Bien que l'IA simplifie effectivement le texte, elle ne semble pas chercher à maximiser systématiquement la facilité de lecture comme le font des experts humains.

Compte tenu de la très forte convergence observée entre les indices de Flesch et de Kandel-Moles dans notre corpus, les analyses qui suivent s'appuieront principalement sur les scores de Kandel-Moles, car cet indice, conçu spécifiquement pour le français, paraît en effet plus adapté à notre projet d'étude. Nous avons néanmoins conservé l'indice de Flesch afin de faciliter la comparaison avec les travaux internationaux qui s'y réfèrent régulièrement.

Au-delà des tendances globales observées au niveau des moyennes, l'analyse des différents extraits permet d'en affiner l'interprétation. Dans la VE, plusieurs scores de Kandel-Moles dépassent 95 : l'extrait VE5 atteint 99,6, tandis que les extraits VE7 et VE10 obtiennent respectivement 98,5 et 98 (voir *Annexe 5*). L'homogénéité est d'ailleurs visible dans le quasi-totalité des extraits de la VE. Contrairement aux autres versions, aucun passage ne s'écarte fortement de la tendance générale, ce qui témoigne d'une stratégie de simplification relativement stable.

Nous notons également que la VIA présente la même tendance générale, avec des scores souvent un peu plus faibles que ceux de la VE. Certains passages atteignent néanmoins des niveaux comparables, notamment VIA14 (97,8) et VIA13 (93,8), qui se situent tous deux à des scores élevés.

Ce point de variabilité est particulièrement intéressant. Alors que les versions éditoriales maintiennent un niveau de lisibilité relativement constant, les textes générés par ChatGPT présentent des écarts plus importants d'un extrait à l'autre. Cela nous laisse penser que les opérations de simplification ne sont pas appliquées avec la même régularité.

Il convient aussi de souligner la forte convergence entre les indices de Flesch et de Kandel-Moles dans notre corpus. Les deux échelles conduisent globalement aux mêmes observations concernant le niveau de difficulté des textes. Les différences relevées entre les versions semblent donc moins liées au choix de l'indice qu'aux stratégies de simplification elles-mêmes. Dans l'ensemble, les versions éditoriales apparaissent plus homogènes, tandis que les productions de l'IA présentent une variabilité plus marquée d'un extrait à l'autre.

3.1.3. Analyse du profil lexical et corrélation avec les indices de lisibilité

Alors que les indices de lisibilité permettent surtout d'évaluer la structure des textes, l'analyse lexicale réalisée avec *Lextutor* apporte un autre éclairage en s'intéressant au vocabulaire utilisé. Afin d'obtenir une vue d'ensemble des tendances observées dans les quatorze extraits du corpus, nous présentons ci-dessous un tableau récapitulatif des valeurs moyennes des principales catégories lexicales pour les trois versions analysées.

Tableau 8. *Moyennes des catégories lexicales selon les trois versions du corpus*

Version	K1	K2	K3	Off-list
VO	80,5%	8,3%	3,5%	7,8%
VE	85,9%	7,3%	2,7%	4,1%
VIA	86,6%	6,6%	3,1%	3,7%

Les données moyennes confirment une augmentation de la proportion de mots très fréquents (K1) dans les deux versions simplifiées, ainsi qu'une réduction importante des mots rares ou spécialisés. À première vue, les deux formes de simplification semblent ainsi produire des effets assez proches.

Dans l'ensemble, les résultats vont dans le même sens que ceux observés avec les indices de lisibilité. Les deux versions simplifiées dépassent le seuil de 85% de mots très fréquents (K1) généralement recommandé pour un public de niveau A2 : 85,9% pour la VE et 86,6% pour la VIA, contre 80,5% dans la VO. Cette proportion relativement élevée de mots K1 dans la version originale peut s'expliquer par le choix volontaire de passages narratifs et dialogués, sélectionnés pour leur potentiel d'adaptation au niveau A2, ainsi que par la segmentation des extraits ne reflétant pas l'ensemble de la complexité stylistique de Zola, plutôt que par une recherche de représentativité exhaustive de cette complexité.

Dans le même temps, la proportion de mots plus rares ou spécialisés (Off-list) diminue nettement, passant de 7,8% dans la VO à 4,1% dans la VE et à 3,7% dans la VIA. Ces chiffres nous donnent d'abord l'impression que les deux formes de simplification produisent des effets assez similaires.

Derrière ces résultats globalement proches, certaines différences apparaissent dans les stratégies de simplification employées. Les versions éditoriales semblent souvent chercher à éliminer les éléments lexicaux susceptibles de poser problème aux apprenants. Plusieurs versions éditoriales illustrent bien cette évolution. Dans les VE7, VE9, VE12 et VE14, les catégories K3 et Off-list ont complètement disparu (voir *Annexe 5*). Les versions générées par ChatGPT aboutissent à des résultats proches, mais elles conservent parfois quelques mots rares, même lorsque le reste du texte est fortement simplifié. Cette particularité passe facilement inaperçue dans les résultats globaux, mais elle ressort lorsqu'on examine chaque extrait séparément.

VIA4 est particulièrement révélateur. Si l'on se fie uniquement au score de Kandel-Moles, ce texte paraît très accessible, avec un score de 80,4. Ce résultat s'explique en grande partie par des phrases très courtes (ASL de 10,9). Pourtant, son profil lexical raconte une toute autre histoire. On y trouve encore 10.3% de mots Off-list et 6,5%

de mots de niveau K3. Fait intéressant, la proportion de mots Off-list est même supérieure à la moyenne observée dans le texte original. Ce contraste montre que la simplification syntaxique et la simplification lexicale n'évoluent pas toujours de la même manière. Un texte peut donc présenter une structure plus simple tout en conservant un vocabulaire relativement exigeant. Nous retrouvons cette situation, dans une moindre mesure, dans certaines versions éditoriales. Ainsi, la VE2 obtient un score élevé au test de Kandel-Moles (83,9), alors que la proportion de mots Off-list atteint encore 7,1 %. Ce maintien d'une partie du vocabulaire d'origine semble traduire la volonté de conserver certains termes caractéristiques de l'univers de Zola, sans doute parce qu'ils contribuent au contenu ou à l'atmosphère du récit.

À l'inverse, plusieurs extraits présentent une cohérence entre simplification syntaxique et celle lexicale. VE9 en offre un bon exemple : son score de Kandel-Moles atteint 96 et l'ensemble du vocabulaire appartient aux catégories les plus fréquentes (100% de K1 et K2). Certaines productions de ChatGPT aboutissent également à ce type de résultat. Dans VIA14, le score de Kandel-Moles est de 97,8, tandis que la proportion de mots K1 atteint 94% et que les mots Off-list ne représentent plus que 0,4% du texte. Dans ces cas-là, les deux indicateurs convergent et suggèrent une accessibilité réellement élevée.

L'observation des différents extraits met en évidence deux logiques de simplification relativement distinctes. Les versions éditoriales privilégient généralement une simplification lexicale plus poussée, avec une volonté visible de réduire au maximum la présence de mots complexes. Les versions générées par ChatGPT obtiennent souvent de très bons scores de lisibilité grâce à une forte simplification syntaxique et à un découpage important des phrases. Cependant, cette stratégie ne s'accompagne pas toujours d'une simplification équivalente du vocabulaire. Pour cette raison, certains textes paraissent faciles selon les indicateurs de lisibilité tout en conservant des difficultés lexicales susceptibles de freiner la compréhension d'un apprenant de niveau A2.

3.2. Analyse des procédés de simplification

Après avoir présenté les résultats des mesures quantitatives, nous nous intéressons maintenant aux transformations observables dans les textes eux-mêmes. Cette analyse permet d'identifier les procédés de simplification privilégiés par les simplificateurs humains et par ChatGPT.

3.2.1. Suppression et découpage syntaxique

Les écarts entre la VE et la VIA apparaissent dès que l'on observe la quantité d'informations conservées dans les textes. Les versions éditoriales se situent clairement dans une logique de réduction forte, parfois très forte, avec des niveaux de suppression allant de 40–70% jusqu'à plus de 85% dans certains cas (voir *Annexe 6*). Cela se traduit par des textes qui ne gardent qu'un noyau narratif très restreint, au détriment de nombreux éléments descriptifs ou contextuels présents dans la version originale.

Dans plusieurs extraits, cette logique est visible presque ligne par ligne. Des détails techniques comme les noms de tissus, des précisions architecturales, ou encore des adjectifs très descriptifs du type « tunique de collégien » (VO6) sont régulièrement éliminés. Même certaines images plus littéraires ne survivent pas à ce processus. Les figures imagées aux « glaciers » ou aux « steppes » (VO4) disparaissent également (voir *Annexe 2*), ce qui donne un texte beaucoup plus direct, mais aussi beaucoup plus plat sur le plan stylistique. Dans plusieurs cas, il ne reste finalement que les informations nécessaires à la compréhension de l'action.

Les versions générées par ChatGPT suivent une autre logique. Dans la plupart des extraits, le taux de suppression reste limité, généralement inférieur à 20 % ou compris entre 20 % et 40 %. Les VIA ne cherchent donc pas à réduire fortement le texte. Elles suppriment plutôt certains éléments jugés difficiles, tout en conservant beaucoup d'autres. Les informations concrètes, comme les lieux ou les perceptions des personnages, sont souvent maintenues, alors que quelques mots rares ou fortement marqués sur le plan stylistique disparaissent, par exemple « vieil or » (VO5) ou « mauvette » (VO6). Dans l'ensemble, les VIA restent moins épurées que les VE et conservent une plus grande partie de la trame narrative.

Le découpage syntaxique montre également des différences entre les deux types de versions. Les VE poussent très loin la fragmentation. Les phrases de Zola sont presque systématiquement cassées en segments courts, souvent construits sur un schéma simple sujet–verbe–complément. Dans la VE2, tout un passage décrivant plusieurs services est réduit à une seule phrase énumérative. Dans la VE12, nous observons une succession d’actions isolées. Cette observation rejoint les résultats présentés dans l’analyse de la longueur moyenne des phrases (ASL), où les VE affichaient généralement des valeurs plus faibles que les VIA. Les experts humains semblent ainsi privilégier des phrases particulièrement courtes afin de maximiser l’accessibilité du texte.

Les VIA découpent aussi, mais d’une autre manière. Les phrases complexes sont séparées, parfois de façon nette, même si les liens logiques sont maintenus. Des connecteurs simples sont ajoutés, des propositions sont transformées en phrases indépendantes, sans forcément casser entièrement la continuité. On obtient donc un texte qui reste simplifié, mais qui « respire » davantage. La progression narrative est moins brutalement interrompue que dans les VE.

Les différences observées au niveau du traitement du contenu se retrouvent dans les données de longueur textuelle. Nous y trouvons les grands écarts. Dans l’extrait 4, on passe de 363 mots pour la VO à seulement 51 mots pour la VE, soit une réduction de plus de 85%, et à 108 mots pour la VIA. L’écart est suffisamment important pour modifier sensiblement la densité informationnelle du texte. Dans l’extrait 7, la logique est similaire : 315 mots pour la VO, 34 pour la VE, 111 pour la VIA. La VE isole un seul fil narratif, presque comme une extraction, tandis que la VIA maintient plusieurs éléments périphériques. Ces différences apparaissent encore plus clairement dans la synthèse chiffrée ci-dessous.

Tableau 9. *Taux de suppression textuelle observés dans les versions simplifiées*

Extrait	VE	VIA
1	57%	18%
2	66%	26%
3	66%	24%
4	86%	70%
5	68%	10%
6	64%	34%
7	89%	65%
8	78%	39%
9	88%	59%
10	78%	62%
11	58%	35%
12	72%	23%
13	53%	28%
14	64%	24%

Alors que les taux de suppression textuelle des VE dépassent fréquemment 60%, voire 80% dans les VE4, VE7 et VE9, les VIA présentent généralement des réductions plus modérées. Même dans les extraits où ChatGPT simplifie fortement le texte, comme les VIA4 (70%) ou VIA7 (65%), le niveau de suppression reste inférieur à celui observé dans les versions éditoriales.

Les résultats présentés ici montrent que les VE et les VIA ne réduisent pas le contenu dans les mêmes proportions. Toutefois, un taux de suppression élevé ne constitue pas en soi une garantie d'adaptation à un niveau A2. Cette question sera abordée plus loin à travers les évaluations des enseignants.

3.2.2. Substitution lexicale et reformulation

Le remplacement du vocabulaire rare ou spécialisé constitue l'un des procédés les plus fréquents dans les deux types de simplification. Dans de nombreux extraits, les

termes anciens, techniques ou peu fréquents disparaissent au profit d'un vocabulaire plus courant, généralement associé aux niveaux les plus fréquents du profil lexical. En regardant les extraits de plus près, les choix effectués ne sont pas toujours les mêmes. Lorsque Zola évoque une « étroite cellule mansardée » (VO1), la version éditoriale conserve seulement l'idée générale du lieu en la remplaçant par « chambre » ou « pièce » (VE1) (voir *Annexe 2*). La version générée par ChatGPT procède souvent autrement. Face à des réalités devenues peu familières pour un lecteur contemporain, il cherche davantage à moderniser le terme qu'à le supprimer complètement. Ainsi, « fenêtre à tabatière » (VO1) devient « petite fenêtre » (VIA1), « redingote » (VO6) est remplacé par « veste » (VIA6) et les « sergents de ville » (VO7) sont transformés en « policiers » (VIA7). Dans ces différents cas, les deux versions réduisent la difficulté lexicale. Mais leurs logiques ne sont pas identiques : l'une tend à généraliser, tandis que l'autre cherche plus souvent un équivalent contemporain.

Nous marquons également cette différence lorsque les textes contiennent des expressions fortement marquées sur le plan littéraire. Les images et les comparaisons utilisées par Zola posent souvent un problème particulier pour un public A2, car leur compréhension suppose non seulement des connaissances lexicales, mais aussi un travail d'interprétation. Les solutions observées dans le corpus sont parfois très éloignées. L'image de « l'infini des steppes » et de « l'entassement des glaciers » (VO4) utilisée par Zola pour comparer l'étendue des piles de linge blanc à un paysage sans fin, disparaît presque entièrement dans la VE pour laisser place à une formulation beaucoup plus directe : « Il y a déjà les fleurs blanches ! » (VE4). Dans la VIA, la référence métaphorique n'est pas conservée telle quelle, mais son sens visuel reste perceptible à travers une reformulation comme « Partout, on voyait du blanc » (VIA4). Nous retrouvons aussi un phénomène comparable avec l'expression « cité du travail » (VO3). Là où certaines VE privilégient la suppression des images complexes, la VIA propose une reformulation explicative : « une petite ville de travail » (VIA3). La même tendance s'observe avec « réclame vivante » (VO2), dont le fonctionnement est expliqué par une description des enfants portant des ballons

publicitaires (VIA2). Les métaphores ne sont donc pas toujours éliminées ; elles tendent parfois à être transformées en descriptions plus transparentes.

Les passages consacrés aux émotions ou aux états psychologiques montrent également des choix différents. Dans plusieurs extraits, les formulations complexes présentes dans le texte original sont réduites à des émotions directement identifiables. Certaines versions éditoriales résument ainsi des développements psychologiques relativement longs par des formulations simples comme « Denise était très émue » (VIA1). La reformulation opérée par ChatGPT conduit souvent au même résultat en termes de compréhension, mais elle conserve davantage d'éléments contextuels autour de l'émotion décrite. Les frontières entre suppression et reformulation ne sont d'ailleurs pas toujours faciles à établir, plusieurs transformations combinant tous ces deux procédés.

Il est intéressant de noter que les termes techniques, historiques ou culturellement marqués donnent lieu à des adaptations similaires. Lorsque Zola utilise le mot « commis » (VO3), la version éditoriales retient une reformulation très générales : « gens qui travaillent » (VE3), tandis que la VIA choisit « employés » (VIA3), terme plus proche du référent initial mais plus accessible pour un lecteur contemporain. Le même constat peut être fait pour certains mots ayant une forte connotation sociale ou morale. Dans plusieurs extraits, la VIA remplace des termes du XIXe siècle par des expressions moins marquées. Ainsi, « gueuse » (VO14) devient « mauvaise femme » (VIA14) et certaines occurrences de « maîtresse » (VO11) sont reformulées en « employées » (VIA11). Cette simplification modifie parfois le sens ou la valeur des mots dans le texte original, mais elle facilite aussi la compréhension lorsque ces termes renvoient à des réalités historiques ou culturelles peu familières pour des apprenants de FLE.

3.2.3. Explicitation

Dans notre corpus, l'explicitation se manifeste à travers plusieurs procédés complémentaires observés dans la VE et la VIA.

Le premier effet observable se situe au niveau de l'identification des personnages et le cadrage de la situation. L'identification des figures subit une accélération notable,

comme le montre la présentation immédiate de Denise en « vendeuse » dans la VE1 (voir *Annexe 2*). Contrairement au texte source qui installe cette information avec le temps, la VE l'annonce tout de suite, modifiant un peu l'image que le lecteur se fait du personnage. Le même phénomène s'observe dans la VE7. Madame Desforges est introduite plus directement, ce qui aide le lecteur à comprendre plus rapidement qui elle est. De la même manière, dans la VIA2, l'expression « réclame vivante » est remplacée par une reformulation qui décrit simplement des enfants portant des ballons publicitaires. Le lecteur n'a donc plus besoin d'interpréter cette image pour comprendre la scène. Ces choix rendent la situation plus explicite et facilitent la compréhension du passage.

Cette tendance se poursuit ensuite dans la structuration des relations logiques. Les liens causaux ou temporels, souvent suggérés mais non entièrement formulés chez Zola, sont ici rendus visibles de manière plus explicite. L'ajout de marqueurs comme « Maintenant » ou « À présent » (VE3) permet ainsi de signaler sans ambiguïté les ruptures ou les évolutions de la situation narrative. Ce même mouvement apparaît aussi dans la VIA5, la relation causale est directement formulée lorsque le statut de Mouret comme chef est expliqué par le fait que « tous les employés le saluaient ». Dans le texte original, le lecteur doit souvent faire lui-même le lien entre les différents éléments. Les versions simplifiées rendent ce travail plus facile. Ce qui restait implicite est plus souvent formulé directement. Les deux versions ne procèdent cependant pas tout à fait de la même manière. Les VE ajoutent plus volontiers des repères qui situent les événements dans le temps. Les VIA expliquent plus directement les liens de cause à effet.

À un autre niveau, il convient de noter la dimension émotionnelle et d'échanges entre les personnages. Dans le texte de Zola, les émotions sont souvent suggérées par de petits indices répartis tout au long du récit. Dans les versions simplifiées, elles sont plus souvent exprimées de manière explicite. Ainsi, Denise est décrite comme « très émue » dans la VIA1. Dans d'autres passages, des sentiments plus nuancés sont résumés par des formulations courtes comme « mal au cœur » (VIA9) ou « Il pleure » (VE14). Nous retrouvons la même tendance dans les dialogues. Des indications

telles que « répond durement » (VIA12) ou « très fâchée » (VIA11) orientent davantage l'interprétation du lecteur. Dans le texte original, ces émotions étaient plutôt suggérées par le contexte et les échanges entre les personnages. Nous observons ici une convergence entre la VE et la VIA dans la verbalisation des affects, même si la VIA cherche à expliciter de manière plus descriptive les intentions interactionnelles. Cela donne l'impression que l'effet global tend vers une transparence accrue des intentions et des affects, même si cette clarification peut parfois lisser la complexité psychologique propre au texte source.

Pour finir, certains ajouts visent à expliciter des implicites d'ordre socio-culturel, susceptibles de poser des difficultés d'interprétation. La VE10 illustre bien cette tendance. La phrase « Tout le monde sait ce que cela veut dire » explicite un sous-entendu qui, dans le texte original, est davantage laissé à l'interprétation du lecteur. L'invitation de Mouret à dîner est ainsi présentée d'emblée comme un rendez-vous amoureux, tel que les autres personnages la comprennent déjà. Cette reformulation limite les ambiguïtés et rend le passage plus facile à comprendre pour des apprenants de niveau A2, surtout lorsque l'interprétation dépend de références culturelles ou de pratiques sociales qui ne leur sont pas forcément familières.

3.3. Analyse de la préservation des caractéristiques littéraires

S'il est clair que les versions simplifiées facilitent la lecture pour un public de niveau A2, notamment grâce à la suppression de certaines informations et à l'explicitation de nombreux implicites, leur impact sur la dimension littéraire du texte mérite également d'être examiné. Cette question se pose particulièrement dans notre corpus consacré à Zola, dont l'écriture accorde une place importante aux descriptions, aux sensations et à l'accumulation de détails. À partir des résultats présentés dans les sections précédentes, nous nous intéressons donc ici aux éléments qui sont conservés, transformés ou atténués dans les différentes versions simplifiées.

3.3.1. Préservation des invariants narratifs

Les simplifications n'affectent pas de la même manière l'organisation du récit. Pourtant, les grandes étapes de l'histoire restent généralement reconnaissables.

Malgré certaines différences dans la manière de raconter les événements, le fil narratif demeure facile à suivre.

C'est dans plusieurs extraits que nous retrouvons, assez clairement, cette stabilité de la trame principale. Denise arrive à Paris, rencontre Mouret, traverse une série d'épreuves : le récit aboutit alors à la scène de reconnaissance et à l'aveu final. Ces étapes, aussi bien que dans les VE que dans les VIA, apparaissent nettement dans les extraits 1, 5, 8, 13 et 14 (voir *Annexe 2*). À ce niveau-là, rien n'est vraiment perdu. Les différences se trouvent surtout dans la façon de relier ces épisodes entre eux. Les principaux événements restent présents dans les deux versions. En revanche, ils ne sont pas racontés de la même manière.

Les deux types de simplification se distinguent particulièrement dans leur manière d'organiser le déroulement du récit. Les VE vont plus rapidement vers l'événement principal. Une partie de ce qui prépare la scène disparaît. Les informations de contexte sont réduites et le lecteur arrive plus vite au moment essentiel. La VE4 en donne un bon exemple. La scène est fortement raccourcie et le geste de Mouret offrant les violettes devient presque le seul élément conservé. On retrouve le même fonctionnement dans les VE7, VE9 et VE12. À chaque fois, le récit se concentre sur un seul moment, tandis que plusieurs étapes qui y conduisaient disparaissent. Les VIA suivent généralement une autre logique. Les événements sont eux aussi simplifiés, mais leur enchaînement reste plus proche de celui du texte original. Les différentes étapes du récit sont plus souvent conservées. La progression de la scène reste donc plus facile à suivre avant d'arriver au moment central. Cette tendance s'observe notamment dans les VIA2, VIA3, VIA6 et VIA11. À l'inverse, certaines VE donnent davantage l'impression de juxtaposer des informations déjà condensées. Les événements racontés restent pourtant les mêmes ; c'est surtout leur organisation qui diffère.

En ce qui concerne les éléments contextuels, le constat est légèrement différent. Certaines VIA gardent des éléments qui ne sont pas strictement nécessaires à la compréhension de l'action principale. Nous retrouvons, par exemple dans les VIA, que certains détails restent présents alors qu'ils ne jouent qu'un rôle secondaire dans

le récit. C'est le cas des enfants et des nourrices dans la VIA7, ou encore de Hutin dans la VIA10. Ces éléments occupent peu de place, mais ils contribuent encore à faire exister le monde qui entoure les personnages. Dans les VE, ils disparaissent plus facilement. Le récit avance plus vite. La lecture devient plus simple. En même temps, une partie du contexte s'efface.

Nous soulignons un autre mécanisme dans les extraits des experts humains, qui rendent souvent certaines informations plus explicites. L'identification des personnages devient ainsi plus directe, par exemple dans la VE5 où Denise n'hésite plus avant de reconnaître Octave Mouret, son nom apparaît immédiatement. La même tendance se retrouve dans la VE12. La solitude de Denise n'est plus seulement suggérée par les réactions des autres personnages. Elle est formulée directement par la phrase « Personne ne l'aime ». À l'inverse, le texte de Zola laisse davantage le lecteur construire lui-même cette impression. On retrouve le même mouvement dans la VIA14 avec « Il pleure ». Une émotion qui occupait plusieurs lignes chez Zola est ramenée à une formulation très brève. Ce type de reformulation contribue à une lecture plus immédiate des personnages. Il modifie également comment les événements sont perçus, qui deviennent plus directs mais moins nuancés.

3.3.2. Préservation des effets sensoriels et descriptifs

L'analyse des différentes versions montre que la simplification ne touche pas uniquement la structure du récit : elle transforme aussi en profondeur le traitement des éléments descriptifs et sensoriels propres à Zola. D'un extrait à l'autre, les perceptions visuelles, les sons, ou encore les impressions physiques et émotionnelles subissent des fortunes diverses. Le tableau suivant synthétise ces tendances générales avant une analyse plus détaillée.

Tableau 10. *Variations des perceptions sensorielles entre VE et VIA*

Sens / Dimension visuelle et sensorielle	Traitement dans les VE	Traitement dans les VIA
Vision	Neutralisation : description purement factuelle des événements	Représentation réaliste : maintien de la richesse visuelle et de l'effet de surprise
Audition	Suppression ou transformation en description d'objets ou de vêtements	Réactivation des sons à l'aide d'un lexique simple mais identique
Atmosphère de foule	Réduction à une action principale du personnage central	Reconstitution d'une atmosphère vivante des groupes et personnages secondaires

Du point de vue visuel, nous trouvons que les VE ont tendance à remplacer souvent les images fortes de Zola par des formulations plus descriptives et fonctionnelles. Le métaphorique s'efface ainsi derrière un registre plus concret, centré sur l'action immédiate. À l'inverse, les VIA gardent parfois cette densité visuelle d'origine. C'est notamment le cas de la VIA4, où l'atmosphère lumineuse et la dynamique d'ensemble de la scène restent palpables malgré la simplification du lexique.

Même constat est remarqué pour la dimension sonore. Dans l'extrait 1 par exemple, la différence entre VE1 et VIA1 est particulièrement nette, et le contraste est même assez frappant. L'expert humain choisit d'effacer une grande partie des effets sonores. L'attention se porte davantage sur les objets eux-mêmes que sur les bruits qui les accompagnent. La scène reste facile à comprendre, mais elle perd une partie de son ambiance. Dans la VIA, plusieurs repères auditifs sont encore présents. Ils sont reformulés, mais restent reconnaissables. C'est le cas de l'expression « bruit de la

soie », qui disparaît dans la version éditoriale. Même discret, ce détail contribue encore à faire entendre l'ambiance du passage.

Cette logique de simplification se retrouve dans les scènes de foule. La VE7 gomme la complexité du décor en regroupant les détails dans de grands ensembles flous, appauvrissant la perception de la scène ; la VIA7, au contraire, s'avère plus fidèle en conservant les arrière-plans et les personnages secondaires, ce qui offre une vision plus complète de la situation. Nous observons d'ailleurs un procédé similaire dans l'extrait 10, où les réactions physiques des personnages, purement et simplement coupées dans la VE, survivent sous une forme allégée dans la VIA.

3.3.3. Le compromis entre l'esthétique narrative zolienne et l'accessibilité du texte

Nos analyses précédentes ont mis en évidence des effets de simplification qui concernent aussi bien la langue que l'organisation du récit, même si ces deux niveaux ne sont pas toujours modifiés de la même manière. Une autre question se pose toutefois, moins évidente : qu'est-ce qui subsiste réellement de l'expérience de lecture proposée par Zola lorsque le texte est adapté à un niveau A2 ?

Dans la VE5, le lecteur apprend presque immédiatement que l'homme observé par Denise est Octave Mouret (voir *Annexe 2*). L'information est rendue rapidement accessible, avec une forme de clarté et de directivité qui laisse peu de place à l'ambiguïté. Pourtant, dans sa VO, l'intérêt de la scène ne tient pas uniquement à cette révélation. Ce qui compte, c'est aussi la progression qui s'y construit. La reconnaissance de Mouret se construit peu à peu, à travers le regard de Denise. Lorsque cette progression narrative disparaît, le lecteur n'accompagne plus vraiment le personnage dans sa découverte et se retrouve davantage face à un résultat déjà constitué, sans avoir suivi les étapes qui y mènent.

Nous retrouvons une logique presque inverse dans la VIA1, où le changement de statut de Denise ne se réduit pas à un constat. Le déplacement dans l'espace, le changement d'habit et plusieurs étapes intermédiaires demeurent présents malgré la simplification. Le lecteur continue malgré tout à suivre un certain processus de construction du sens, même si celui-ci est partiellement simplifié. Ce n'est donc pas seulement l'installation de Denise dans sa nouvelle chambre et le moment où elle

enfile son uniforme de travail qui est raconté, mais aussi le passage progressif de son état d'émotion et de découverte vers un état de malaise et de honte. La lecture demande un peu plus d'attention, mais elle reste globalement accessible à un lecteur de niveau A2.

En lisant la VE9, on est surtout frappé par la place accordée à la souffrance de Mouret. Le grand chantier passe presque au second plan. Pourtant, dans le roman de Zola, il ne sert pas seulement de décor. Il rappelle sans cesse le contraste entre la réussite commerciale de Mouret et le vide affectif dans lequel il se trouve. La VIA9 ne reprend pas tous les détails du texte original. Malgré cela, cette opposition reste encore perceptible. Le lecteur retrouve à la fois l'ampleur du chantier et la solitude du personnage. Cette idée, importante dans le roman, continue ainsi d'accompagner la lecture : la réussite matérielle suffit-elle vraiment lorsque le bonheur personnel fait défaut ?

Le même phénomène apparaît dans la VIA13, mais sous une autre forme. Lorsque Denise rappelle qu'elle travaille depuis son enfance, cette précision peut sembler discrète. En réalité, elle change la manière de comprendre son refus. Celui-ci ne relève plus seulement d'un choix personnel. Il renvoie aussi à son rapport au travail, à l'indépendance qu'elle a construite au fil des années et à la valeur qu'elle accorde à son propre parcours. Dans la VE, cet arrière-plan devient beaucoup moins visible. Le refus de Denise est toujours présent. Ce qui le nourrit l'est beaucoup moins. Ce déplacement est bien intéressant, il touche directement à l'image du personnage. D'un côté, nous avons surtout un événement ; de l'autre, nous retrouvons encore une partie des principes qui le motivent.

Cette présence de justifications ne concerne pas uniquement les motivations individuelles des personnages. Nous notons des moments où la simplification ne touche plus seulement ce qui est raconté, mais la manière dont la scène est perçue. Dans la VIA12 par exemple, les rires ne viennent plus d'un groupe anonyme. Ici, Marguerite et Clara sont nommées. La scène prend alors une autre dimension. Denise semble davantage isolée, parce que le lecteur voit plus clairement de qui viennent les réactions. Le rejet n'est jamais formulé de manière explicite. Pourtant, il se ressent

au fil de la lecture. La VE fait un choix différent. Une partie de ces éléments disparaît et le conflit est présenté de façon plus directe. La scène reste facile à comprendre, mais elle laisse moins de place à ce qui entoure la situation. En comparant les deux versions, nous avons eu l'impression que la position du lecteur changeait elle aussi. Dans la VE, le conflit paraît déjà installé. Avec la VIA, on perçoit encore les petites réactions qui conduisent progressivement à cette mise à l'écart. Cette nuance change légèrement la manière dont on traverse le texte, sans pour autant en modifier les faits principaux.

La fin du roman va dans le même sens. Le mariage a toujours lieu, quelle que soit la version. Pourtant, lorsque les hésitations de Denise sont conservées — sa peur des commérages, son refus d'être perçue comme une « mauvaise femme » (VIA14) parce qu'elle épouse son patron — la signification du dénouement se nuance. L'union n'apparaît plus comme une simple récompense sentimentale. Les tensions sociales qui entourent cette ascension restent visibles et rappellent que la réussite de Denise demeure, jusqu'au bout, un sujet de jugement et de discussion.

Ces observations rejoignent d'ailleurs les résultats obtenus concernant la charge cognitive (voir *Annexe 7*). Les VE réduisent fortement l'effort interprétatif demandé au lecteur : les informations essentielles sont explicitées, les liens entre les événements sont souvent rendus plus transparents et la progression narrative devient plus facile à suivre. Les VIA demandent un peu plus d'efforts au lecteur. Tout n'est pas expliqué de manière explicite. Il faut parfois faire le lien entre plusieurs informations, suivre l'évolution des personnages ou comprendre une relation de cause à effet qui reste seulement suggérée. Pour un apprenant de niveau A2, cette lecture reste néanmoins possible. En contrepartie, une partie du mouvement du récit et de l'atmosphère du texte original est davantage préservée.

Les différences observées ne racontent pourtant pas une histoire aussi simple. Elles ne permettent pas de dire qu'un type de simplification fonctionne toujours mieux que l'autre. Chaque extrait produit des effets un peu différents. Les choix de réécriture jouent aussi un rôle. Au fil de l'analyse, nous avons surtout constaté que la

simplification modifie la manière dont le texte est reçu. Ce que le lecteur comprend change. Ce qu'il remarque aussi.

La réflexion dépasse alors les seuls aspects linguistiques. Ce qui est en jeu ne s'agit pas seulement du texte lui-même, mais aussi la façon dont il sera lu en classe. Chez des apprenants de niveau A2, la lecture est rarement entièrement autonome. Elle est souvent accompagnée par l'enseignant, expliquée de nouveau lorsque cela est nécessaire et parfois enrichie par les échanges qui ont lieu pendant l'activité. C'est précisément à partir de ces écarts que se pose la question de la médiation pédagogique. Ce point sera approfondi dans la section suivante consacrée à la perception des enseignants face à ces différentes formes de simplification.

3.4. Analyse des résultats expérimentaux : le regard des enseignants du FLE

3.4.1. Profil des participants

Notre enquête en ligne adressée à des enseignants du FLE, diffusée au début du mois d'avril 2026, nous a permis de recueillir 28 réponses exploitables, constituant ainsi notre échantillon d'analyse.

Nous voudrions commencer par présenter brièvement le profil des enseignants ayant participé à l'enquête, car, selon nous, certaines caractéristiques, le contexte d'enseignement ou l'expérience professionnelle par exemple, peuvent influencer leur perception de la simplification des textes.

Dans cet échantillon, la répartition entre locuteurs natifs (53,6%) et non natifs (46,4%) est presque équilibrée (fig. 3). Nous avons inclu cette question dès le début du questionnaire parce qu'une interrogation assez simple s'est posée lors de la préparation de l'enquête : un enseignant ayant lui-même appris le français comme langue étrangère porte-t-il le même regard sur un texte simplifié qu'un locuteur natif ? Cette variable nous permettra donc d'explorer cette question dans les sections suivantes.

Êtes-vous locuteur natif du français ?

28 réponses

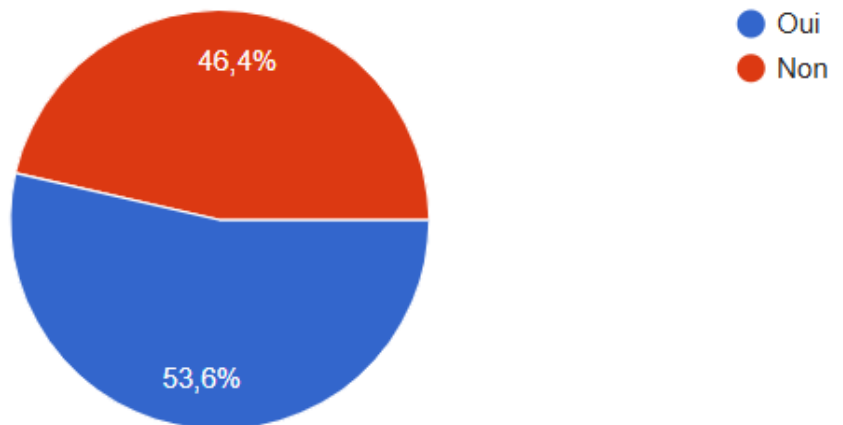


Figure 4. Répartition des réponses obtenues à notre questionnaire (avril 2026)

Les participants disposent également d'une expérience professionnelle solide. Tous ont déjà enseigné le FLE et plus de huit sur dix (82,1%) sont encore en activité aujourd'hui (fig. 4). Si près de la moitié des enseignants (46,4%) justifient de plus de dix ans d'expérience (fig. 5), la proportion restante répartit entre les profils intermédiaire (32,1% ayant entre cinq à dix ans d'ancienneté) et les enseignants plus récemment entrés dans la profession (21,4%). L'échantillon rassemble ainsi des enseignants disposant d'une expérience suffisamment importante pour porter un jugement réfléchi sur les supports pédagogiques proposés.

Avez-vous déjà enseigné le FLE ?

28 réponses

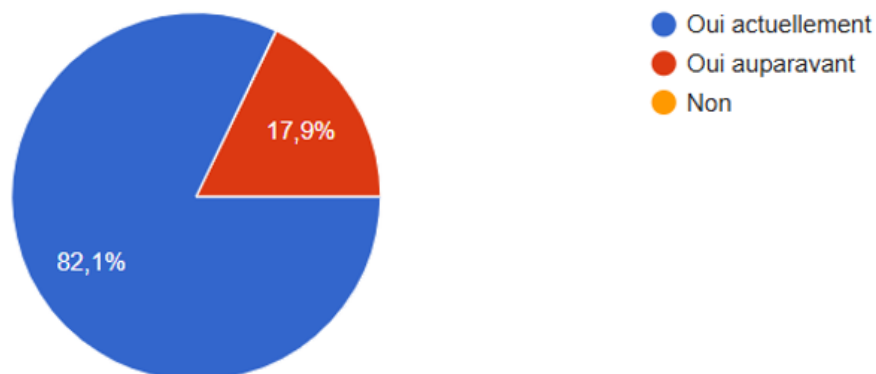


Figure 5. *Enquête, suite (pratique du FLE)*

Nombre approximatif d'années d'expérience en enseignement

28 réponses

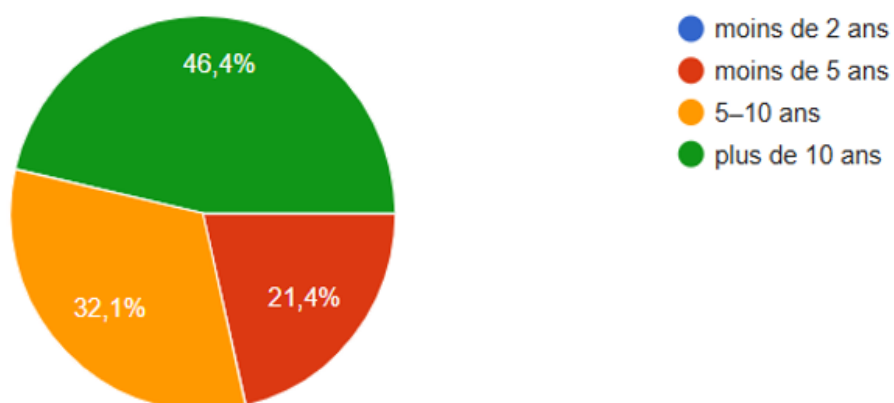


Figure 6. *Enquête, suite (ancienneté)*

Les contextes d'enseignement représentés sont, eux aussi, relativement diversifiés. Une légère majorité des répondants exerce dans des pays non francophones (53,6%), contre 46,4% dans des espaces où le français est langue officielle (fig. 6). Rien d'étonnant ici, mais cela nous rappelle que le rapport à la langue dépend avant tout de l'environnement professionnel.

Si vous avez déjà enseigné le FLE, est-ce dans un pays dont l'une des langues officielles est le français ?

28 réponses

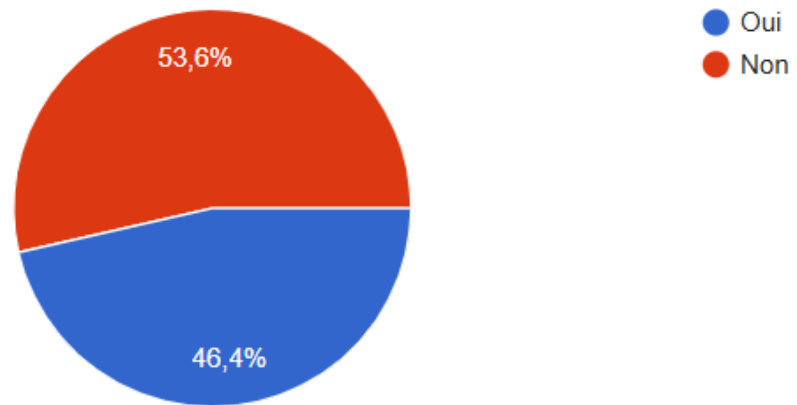


Figure 7. *Enquête, suite (contexte)*

3.4.2. Évaluation globale des extraits simplifiés

À partir des résultats ci-dessus, nous proposons à présent une lecture plus globale des évaluations des extraits simplifiés. Ce qui nous surprend, c'est que les résultats montrent d'abord une hétérogénéité dans les jugements. Aucune tendance ne reste stable ici.

Quatre dimensions qui structurent les évaluations des enseignants sont la clarté pour un niveau A2, le potentiel pédagogique, l'interaction en classe et la capacité d'intégration dans un dispositif d'enseignement. Mais dans les faits, ces quatre axes ne fonctionnent jamais séparément. Ils se croisent, se contredisent parfois, et finissent par produire des hiérarchies inattendues entre la VE et la VIA.

Avant d'interpréter ces résultats, il est important de garder un élément en tête. Au moment de l'évaluation, les enseignants ne savaient jamais s'ils lisaient une version éditoriale ou une version générée par ChatGPT. Les extraits avaient été mélangés et n'étaient pas présentés dans l'ordre. Les extraits E6, E8 et E11 apparaissaient chacun sous leurs deux versions simplifiées, sans aucun indice permettant de les distinguer. Les enseignants n'avaient donc qu'une seule chose sous les yeux : le texte lui-même.

La clarté ne suit pas une logique stable d'un extrait à l'autre, et ce qui ressort surtout, c'est l'impossibilité de dégager une tendance unique entre les VE et les VIA sur ce critère. Dans certains cas, les VE semblent mieux évaluées, mais cet avantage disparaît immédiatement dans d'autres configurations, tandis que certaines situations aboutissent simplement à une égalité. C'est ce que nous montrent concrètement les données : la VE est en tête sur E6 avec 3,8/5 contre 3,4 pour la VIA, mais la tendance s'inverse sur E8 où la VIA atteint 3,5 face à 2,8 pour la VE, avant que E11 ne neutralise complètement la comparaison avec un score identique de 3,3 pour les deux versions. Nous présentons ensuite les résultats détaillés dans le tableau ci-dessous.

Tableau 11. *Évaluation des extraits selon les quatre critères*

Extrait		Intégration	Clarté	Potentiel	Interaction
E6	VE	2.6	3.8	2.9	2.5
	VIA	3.1	3.3	3.5	3.6
E8	VE	2.9	2.8	3.2	3.3
	VIA	3	3.5	3.4	3.1
E11	VE	2.9	3.3	2.8	2.9
	VIA	2.4	3.3	3.5	3.6

Si la clarté déroute, le potentiel pédagogique offre, lui, une trajectoire beaucoup plus lisible. Dans les trois cas observés, la VIA obtient des scores supérieurs ou équivalents à la VE, avec un avantage net dans certains cas, notamment 3,5 pour la VIA11 contre 2,8 pour le VE11. Cette tendance se répète de manière assez régulière. Ce résultat ne semble pas lié uniquement à des aspects techniques du texte. Il indique plutôt que les enseignants ne se limitent pas à juger la facilité de compréhension. Ils prennent aussi en compte l'usage possible du texte en classe, ce que l'on peut en faire dans un contexte pédagogique, les activités qu'il peut permettre ou les échanges qu'il peut générer.

Et cette idée devient encore plus visible lorsqu'on regarde le critère de l'interaction en classe. Pour les extraits E6 et E11, les VIA obtiennent des scores plus élevés sur ce critère. L'écart est même assez net : 3,6 contre 2,5 pour l'E6, puis 3,6 contre 2,9

pour l'E11. Ces résultats ont bien attiré l'attention. Ils ne signifient pas pour autant que les VIA sont plus efficaces simplement parce qu'elles sont moins ou plus simplifiées. Il reste encore d'autres éléments qui entrent certainement en jeu. Les enseignants évaluent aussi la manière dont un texte peut être utilisé en classe, les échanges qu'il permet de susciter ou les activités qu'il rend possibles. Les notes obtenues par les VIA montrent donc qu'elles sont souvent perçues comme plus intéressantes sur ce point, sans que la simplification puisse, à elle seule, expliquer cet écart.

Pour l'intégration en classe, aucune régularité ne se dégage. E6 et E8 favorisent la VIA. E11 fait exactement l'inverse, la VE passe devant (2,9 contre 2,4 pour la VIA). Dans les données, la notion de support exploitable varie fortement selon le cas. Cette notion dépend des caractéristiques intrinsèques du texte, telles que sa densité et sa lisibilité.

3.4.3. Comparaison des perceptions des enseignants selon le type de simplification

Les analyses précédentes ont mis en évidence des écarts parfois importants d'un extrait à l'autre. Cette diversité des évaluations nous a amenée à regarder les résultats sous un autre angle. Plutôt que de comparer chaque texte séparément, il nous paraît intéressant d'observer ce qui se passe lorsque les extraits sont regroupés selon leur type de simplification.

Afin d'obtenir une vue d'ensemble, nous avons calculé pour chaque extrait une moyenne générale à partir des quatre critères évalués par les enseignants. Le tableau suivant présente le classement des extraits, du score moyen le plus élevé au plus faible.

Tableau 12. *Classement des extraits selon les scores globaux des enseignants*

Rang	Extrait	Score global
1	E6 (VIA)	3.4
2	E8 (VIA)	3.3
3	E11 (VIA)	3.2
4	E8 (VE)	3.1
5	E6 (VE)	3
	E11 (VE)	3

Le classement général apporte déjà un premier élément de réflexion. En tête du classement, nous retrouvons trois versions issues de l'IA : E6, E8 et E11. Les versions éditoriales suivent avec des moyennes légèrement plus faibles. Nous ne nous attendions pas nécessairement à un tel résultat au départ. Les analyses de lisibilité présentées plus haut auraient plutôt pu laisser penser que les textes éditoriaux seraient systématiquement favorisés.

Le tableau 13 suivant nous permet d'aller un peu plus loin en comparant les moyennes obtenues par les deux groupes de textes.

Tableau 13. *Comparaison des moyennes générales entre la VE et la VIA*

Critère	VE	VIA
Intégration en classe	2.8	2.8
Clarté A2	3.3	3.4
Potentiel pédagogique	3	3.5
Interaction en classe	2.9	3.4

Les résultats donnent une impression de proximité entre les deux groupes. L'intégration en classe reçoit exactement la même moyenne dans les deux cas (2,8) et la clarté ne présente qu'un écart minime de 0,1 point. Rien ne permet donc de distinguer clairement un type de simplification de l'autre sur ces dimensions. Le tableau 13 nous révèle néanmoins une configuration différente dès que l'on s'intéresse au potentiel pédagogique et à l'interaction en classe des versions simplifiées étudiées. C'est sur ces deux critères, qui concernent moins la compréhension immédiate du texte que ce qu'il permet de faire une fois en classe, que les écarts sont les plus visibles : les VIA y obtiennent une moyenne supérieure de 0,5 point à celle des VE. Derrière toutes ces évaluations, nous pouvons supposer que les enseignants se projettent déjà dans une situation pédagogique concrète : lancer une discussion, prolonger la lecture par une activité ou encore mobiliser les connaissances des apprenants. Les versions issues de ChatGPT semblent offrir davantage de prises pour ce type d'exploitation.

Le fait que les VIA obtiennent de meilleurs résultats sur ces critères nous a d'ailleurs quelque peu surprise. Nos analyses mettaient déjà en évidence plusieurs caractéristiques des VE susceptibles de faciliter la lecture : des phrases plus courtes, une syntaxe plus simple et une densité lexicale réduite. Les différents indicateurs obtenus conduisaient au même constat.

Les commentaires des enseignants sont, dans l'ensemble, cohérents avec ces résultats. Les différences restent modestes, mais les VE sont le plus souvent considérées comme suffisamment claires pour un public de niveau A2. Autrement dit, les enseignants perçoivent bien les effets de certaines simplifications linguistiques observées dans le corpus.

Le paradoxe apparaît ailleurs. Les textes les mieux classés ne sont pas nécessairement ceux qui obtiennent les meilleurs résultats en matière de lisibilité. Dans plusieurs VIA, certains éléments narratifs, descriptifs ou culturels ont été davantage conservés. Le texte devient parfois un peu plus dense, mais aussi plus riche en contenu. À la lecture des commentaires, nous avons souvent eu l'impression que les participants évaluent simultanément deux choses : la facilité de lecture du support et les possibilités pédagogiques qu'il pouvait ouvrir.

Les premières places occupées par VIA6, VIA8 et VIA11 renforcent cette interprétation. Aucune de ces versions ne correspond exactement au profil du texte le plus simplifié du corpus. Pourtant, ce sont elles qui reçoivent les évaluations globales les plus favorables. On peut également relever une légère incohérence intéressante : la VIA6, qui obtient la meilleure moyenne d'évaluation des enseignants, présente toutefois, selon l'indice de Kandel-Moles, le score de lisibilité le plus faible de ces trois extraits.

La lisibilité reste donc un facteur important, mais elle n'explique pas à elle seule les préférences exprimées par les enseignants. Les VE apparaissent souvent plus contrôlées sur le plan linguistique. Les VIA, de leur côté, semblent mieux préserver certains éléments susceptibles d'alimenter les échanges ou les activités en classe. Les résultats obtenus ici invitent finalement à envisager la simplification comme un

équilibre entre plusieurs exigences plutôt que comme une simple réduction de la complexité linguistique.

3.4.4. Influence du profil des enseignants sur les évaluations

Afin d'examiner l'influence du profil des enseignants sur leurs évaluations, nous avons comparé les moyennes attribuées aux VE et aux VIA selon trois variables (statut linguistique, expérience professionnelle et contexte d'enseignement), dont les détails figurent dans les tableaux suivants.

Tableau 14. *Scores moyens des versions selon le statut linguistique des enseignants*

	Enseignant natif		Enseignant non natif	
	VE	VIA	VE	VIA
Intégration en classe	2,4	2,6	3	3,6
Clarté A2	3,4	2,9	4	3,5
Potentiel pédagogique	2,6	3,1	2,9	3,6
Interaction en classe	2,4	2,9	2,6	3,7

Au moment de construire le questionnaire, nous pensions que la distinction entre locuteurs natifs et non natifs produirait des écarts plus marqués. Les résultats que nous avons obtenus vont dans une autre direction.

Dans l'ensemble, les enseignants non natifs donnent des notes un peu plus élevées que les enseignants natifs, quelle que soit la version évaluée. Cette différence est surtout visible pour les VIA. Le potentiel pédagogique atteint par exemple 3,6 chez les répondants non natifs, alors que les enseignants natifs restent généralement plus prudents dans leurs évaluations.

Malgré cette différence, les deux groupes se rejoignent sur un point. La version éditoriale est, dans les deux cas, celle qui obtient les meilleures notes pour la clarté linguistique. Les natifs lui attribuent 3,4 contre 2,9 à la VIA ; les non natifs 4 contre 3,5. Cette convergence n'est finalement pas très surprenante, car elle rejoint ce que montraient déjà les analyses de lisibilité réalisées dans la section précédente : la VE

est perçue comme la plus facile d'accès, indépendamment du profil linguistique des répondants.

La hiérarchie change dès que l'on s'intéresse à l'utilisation du texte en classe. Les VIA recueillent alors les évaluations les plus favorables. Les enseignants semblent y voir davantage de possibilités pour construire des activités, susciter des échanges ou prolonger la lecture. Les écarts observés sont loin d'être anecdotiques, notamment pour le potentiel d'interaction.

Tableau 15. *Scores moyens des versions selon l'ancienneté des enseignants*

	Moins de 5 ans		5 – 10 ans		Plus de 10 ans	
	VE	VIA	VE	VIA	VE	VIA
Intégration en classe	2,7	2,9	2,6	3,1	2,6	2,9
Clarté A2	3,6	3,1	3,8	3,3	3,3	3,1
Potentiel pédagogique	2,8	3,3	2,7	3,4	2,8	3,3
Interaction en classe	2,6	3,2	2,5	3,4	2,5	3,2

L'ancienneté dans l'enseignement fait apparaître des nuances plus discrètes. Le groupe ayant entre cinq à dix ans d'expérience se distingue par les évaluations les plus favorables à la VIA. Les scores atteignent 3,4 pour le potentiel pédagogique ainsi que pour le potentiel d'interaction. Ce même groupe attribue également la meilleure note de clarté à la VE (3,8). Les enseignants concernés ont déjà plusieurs années de pratique derrière eux, tout en restant relativement proches des réalités quotidiennes de la classe. Leurs évaluations donnent parfois l'impression qu'ils identifient assez clairement les fonctions respectives des deux versions : la VE pour faciliter l'entrée dans le texte et la VIA pour exploiter davantage le texte une fois la compréhension installée.

Les enseignants comptant plus de dix années d'expérience distribuent des notes un peu plus basses dans l'ensemble. Cela ne les conduit pas pour autant à rejeter les versions produites par ChatGPT. Le potentiel d'interaction de la VIA atteint encore 3,2 contre 2,5 pour la VE. Leur regard paraît simplement plus prudent.

Les profils les plus jeunes et les plus expérimentés se rapprochent finalement davantage que prévu. Les écarts restent négligeables et ne modifient pas les tendances observées dans le reste du corpus.

Tableau 16. *Scores moyens des versions selon le contexte d'enseignement*

	Contexte francophone		Contexte non francophone	
	VE	VIA	VE	VIA
Intégration en classe	2,7	3	2,7	3
Clarté A2	3,8	3,2	3,7	3,2
Potentiel pédagogique	2,8	3,4	2,8	3,4
Interaction en classe	2,6	3,4	2,5	3,3

Contrairement à nos attentes, le contexte d'enseignement (dans un pays où le français est une des langues principales ou dans un pays où il ne l'est pas) influence peu la nature des évaluations. Les écarts ici restent presque invisibles. Les différences observées paraissent moins liées au contexte d'enseignement qu'aux qualités attribuées à chaque version : la VE pour sa clarté, la VIA pour ce qu'elle permet de faire en classe.

À force de comparer les deux versions simplifiées, on finirait presque par oublier le niveau général des évaluations. Les moyennes observées demeurent relativement modérées et se situent le plus souvent entre 2,5 et 3,5 sur une échelle de 5, ce qui pourrait être interprété par le fait que les enseignants ne considèrent donc ni les VE ni les VIA comme des supports parfaits. Les écarts observés entre les deux versions apparaissent néanmoins suffisamment marqués pour faire émerger des préférences selon les critères évalués. L'intérêt de ces résultats réside ainsi moins dans le niveau absolu des notes que dans la manière dont les enseignants répartissent leurs évaluations entre clarté linguistique, potentiel pédagogique et capacité à susciter des interactions.

3.4.5. Analyse qualitative des commentaires ouverts

Les commentaires laissés par les enseignants nous apportent un éclairage complémentaire aux résultats quantitatifs. Nous observons que certains éléments apparaissaient déjà dans les moyennes obtenues. Les réponses ouvertes permettent de comprendre plus précisément ce qui a motivé certaines évaluations.

Parmi l'ensemble des extraits, nous trouvons deux textes qui reviennent plus souvent que les autres quand il est question de manque de naturel ou d'inadéquation au niveau A2. Il nous faut rappeler que ces extraits, présentés aux enseignants sous les appellations E1 et E5 lors de l'enquête, correspondent en réalité à la VIA8 et à la VE8 dans le corpus. Afin de faciliter la lecture de cette analyse, nous utiliserons désormais les références exactes des extraits, à savoir VIA8 et VE8.

Les remarques laissées au sujet de la VIA8 concernent avant tout la langue. Les enseignants ne remettent pas vraiment en cause la correction du texte. Ce qui semble les gêner est ailleurs. Plusieurs parlent d'un vocabulaire qu'ils jugent peu varié ou trop répétitif. D'autres disent avoir retrouvé les mêmes formulations à plusieurs reprises. La lecture leur paraît alors moins fluide et perd peu à peu de son rythme. Certains commentaires s'arrêtent également sur l'usage des temps du passé, en particulier du passé composé, qui « ne 'sonne' pas très juste » dans ce contexte narratif. Ce ne sont pas nécessairement des erreurs qui sont relevées, mais plutôt une impression générale de texte produit par l'IA, où les choix linguistiques semblent parfois manquer de souplesse.

La VE8 suscite un autre type de réaction. Cette fois, les commentaires concernent moins la qualité de la langue que le niveau de simplification adopté. Plusieurs enseignants estiment que l'adaptation fait disparaître une partie de la richesse du texte littéraire. Selon eux, l'atmosphère est moins perceptible et certaines nuances émotionnelles sont atténuées. Ce résultat nous a surpris, car nous nous attendions plutôt à voir ce type de remarques à propos de la version générée par ChatGPT. Les réponses nous montrent pourtant une situation un peu différente. Une adaptation éditoriale peut, elle aussi, susciter des réserves lorsque la simplification devient très importante. La VE8 provient d'une adaptation réalisée dans un cadre éditorial

professionnel. Son origine ne suffit donc pas à garantir une meilleure réception auprès des enseignants. À la lecture des commentaires, ce sont surtout les choix de simplification qui sont questionnés. Pour plusieurs répondants, le texte s'éloigne du compromis recherché entre accessibilité et dimension littéraire. Les avis ne vont toutefois pas tous dans le même sens. Le même extrait est également cité parmi ceux dont la longueur paraît la plus adaptée à un public A2. Ce qui est perçu comme une réduction excessive par certains devient, pour d'autres, une qualité. Cette divergence rappelle qu'un texte simplifié n'est jamais évalué uniquement à partir de critères linguistiques. Les attentes liées à la littérature, à l'enseignement et au niveau visé entrent elles aussi dans l'appréciation des enseignants.

En relisant l'ensemble des commentaires, plusieurs idées reviennent d'un extrait à l'autre. Le lexique occupe une place importante. Dans certains cas, quelques textes utilisent des mots trop simples ou trop répétitif. Dans d'autres, les enseignants signalent au contraire la présence de termes spécialisés tels que « mousselines », « soies » ou « draperies », perçus comme éloignés du vocabulaire habituellement maîtrisé par un apprenant A2.

Les versions générées par ChatGPT donnent lieu à d'autres remarques. Plusieurs enseignants disent avoir eu du mal à suivre certains enchaînements dans le récit. Les changements de temps attirent aussi leur attention. Quelques répondants parlent enfin d'un lien qui leur paraît parfois moins clair entre deux phrases. Rien de tout cela ne revient dans tous les commentaires, mais ces remarques apparaissent suffisamment souvent pour être relevées.

La longueur des textes est également discutée. Si certains extraits sont jugés trop longs pour un public débutant, d'autres, au contraire, semblent avoir été raccourcis au point de perdre une partie de leur intérêt. Les réponses montrent surtout qu'il est difficile de définir une longueur qui conviendrait à toutes les situations d'enseignement. Les attentes varient d'un enseignant à l'autre.

Une dernière impression ressort de plusieurs commentaires. Lorsqu'un texte est perçu comme artificiel, ce n'est pas à cause d'un seul élément. Les enseignants expliquent qu'ils ne repèrent pas un texte généré par l'IA à partir d'un seul élément. C'est plutôt

l'accumulation de plusieurs indices qui attire leur attention : des dialogues qui semblent moins naturels, des formulations parfois répétitives ou encore une impression d'écriture trop régulière. Certains décrivent même certains passages comme « mécaniques ». À l'inverse, ils rappellent aussi qu'une simplification trop poussée réalisée par un expert humain peut produire le même effet, comme l'illustre la VE8. Lorsque trop d'éléments disparaissent, le texte cesse d'être perçu comme un texte littéraire à part entière.

Les commentaires concernant l'utilisation des VE sont, dans l'ensemble, assez positifs. Beaucoup d'enseignants disent qu'ils seraient prêts à les intégrer dans leurs cours. Selon eux, ces versions simplifiées permettent d'aborder la littérature sans décourager les apprenants dès les premières pages. Plusieurs enseignants parlent aussi de leur intérêt pour enrichir le vocabulaire, faire découvrir certains aspects de la culture ou préparer progressivement la lecture de l'œuvre originale.

Les commentaires ne sont cependant pas tous enthousiastes. Un répondant nous rappelle que l'univers de Zola reste exigeant pour des apprenants de niveau A2, même après un travail de simplification. D'après lui, certaines situations ou certains thèmes peuvent encore paraître assez éloignés de ce que les apprenants connaissent ou vivent au quotidien. Quelques commentaires évoquent également des problèmes de fluidité susceptibles d'entraîner des malentendus lors de la lecture.

La réaction face à l'utilisation des versions générées par l'IA sont plus partagées. Une partie des répondants se montre ouverte à leur utilisation, notamment parce qu'elles permettent de produire rapidement des supports variés ou adaptés à différents besoins pédagogiques. D'autres expriment davantage de prudence et déclarent ne pas souhaiter utiliser directement ces textes dans leur état actuel.

Une idée revient avec insistance dans les commentaires : la nécessité d'une validation humaine. Plusieurs enseignants acceptent le recours à l'IA à condition que les textes soient relus, corrigés ou adaptés avant leur utilisation en classe. Certains soulignent également que l'origine artificielle du texte soit explicitement indiquée.

Les limites mentionnées concernent principalement le style. Les enseignants évoquent parfois un manque de profondeur émotionnelle, une uniformisation des

formulations ou une impression que le texte « ne sonne pas toujours très français ». Ces remarques n'aboutissent pas à un rejet général des VIA, mais elles montrent que leur acceptabilité pédagogique reste étroitement liée à l'intervention d'un enseignant capable de les contextualiser et de les ajuster.

À la lecture de ces commentaires, ce qui frappe n'est pas tant l'existence d'un consensus que la diversité des attentes exprimées par les enseignants. Certains accordent une importance particulière à la fidélité au texte littéraire, d'autres semblent davantage préoccupés par l'accessibilité pour les apprenants. Quelques remarques paraissent même difficiles à concilier entre elles. Cette variété de points de vue rappelle qu'un texte adapté au niveau A2 ne répond pas à une définition unique et que son appréciation dépend souvent de ce que chaque enseignant considère comme prioritaire dans sa pratique.

Nous trouvons enfin une remarque, qui mérite d'être mentionnée, d'un enseignant qui refuse d'utiliser tant les VE que les VIA, estimant disposer déjà d'autres ressources mieux adaptées au niveau A2. Cet avis demeure marginal dans le corpus. Il montre néanmoins que, pour certains répondants, le choix ne se joue pas uniquement entre une version éditoriale et une version générée par l'IA. La pertinence même de ces adaptations peut également être interrogée au regard des autres supports disponibles pour l'enseignement du FLE.

3.5. Discussion et recommandation

3.5.1. Synthèse et mise en perspective des résultats

Au moment de formuler nos questions de recherche, nous nous attendions à observer une opposition relativement classique entre la simplification humaine et celle artificielle. En avançant dans les analyses, nous avons progressivement constaté que cette opposition de départ tenait moins bien qu'attendu. Certains résultats vont même dans une direction que nous n'avions pas envisagée au début de ce travail.

La première question cherchait à identifier les principales différences morphosyntaxiques et stylistiques entre les VE et les VIA. En confrontant les extraits du corpus, nous avons remarqué une distinction particulièrement nette. Les deux type de simplification poursuivent le même objectif général — rendre le texte

accessible à un public A2 — mais ils n’empruntent pas le même chemin pour y parvenir.

Du côté des VE, c’est par une fragmentation massive qu’elles opèrent : avec une moyenne de 9,3 mots par phrase — le seuil le plus bas du corpus —, les longues périodes de Zola se trouvent littéralement éclatées en unités minimales. Nous observons le même mouvement dans les choix lexicaux. Les termes spécialisés, les descriptions détaillées et certaines formulations imagées sont souvent remplacés par des expressions plus générales. La lecture devient alors plus simple, mais le texte perd parfois une partie de sa richesse.

Les VIA donnent une impression un peu différente. Même si leur longueur moyenne reste faible (10,3 mots), la réduction est généralement moins importante. Les phrases sont raccourcies, mais elles ne sont pas systématiquement découpées. Les connecteurs les plus simples sont souvent conservés, ce qui permet de maintenir les relations logiques entre les idées. La simplification lexicale suit la même logique. Là où la VE tend à neutraliser certains éléments du texte, l’IA privilégie plus souvent la reformulation ou l’actualisation. Le référent change peu ; la formulation devient simplement plus proche de l’usage contemporain.

Ces observations conduisent déjà à nuancer notre deuxième hypothèse. Nous pensions initialement que les VE chercheraient davantage à préserver la richesse lexicale de l’œuvre tandis que l’IA privilégierait une simplification lexicale importante. Les résultats montrent plutôt l’inverse. Dans de nombreux cas, ce sont les VE qui procèdent aux réductions les plus fortes, alors que ChatGPT conserve une partie du vocabulaire, des détails descriptifs ou des références contextuelles présentes dans le texte source.

Dans les réponses des vingt-huit enseignants, nous ne voyons pas de vraie opposition entre la VE et la VIA concernant leur usage en classe de FLE pour travailler des textes littéraires. Dans l’ensemble, les enseignants considèrent les VE comme plus accessibles. Cette perception va dans le même sens que les résultats obtenus lors de l’analyse de la complexité linguistique. Elle ne suffit cependant pas à confirmer nos interprétations et doit être mise en relation avec les autres résultats de l’étude.

Lorsque nous passons des critères de lisibilité aux usages possibles en classe, l'image devient un peu différente. Les VIA obtiennent des scores plus élevés en potentiel pédagogique (3,5 contre 3) et en potentiel d'interaction (3,4 contre 2,9). Les trois extraits arrivés en tête du classement général appartiennent d'ailleurs au groupe des VIA. Cette préférence ne peut pas être expliquée directement par le questionnaire lui-même, mais elle semble pouvoir être rapprochée de plusieurs observations effectuées lors de l'analyse qualitative. Les textes produits par ChatGPT conservent souvent davantage de personnages secondaires, d'éléments contextuels ou de détails liés à l'univers du roman. Cette matière supplémentaire offre probablement davantage de points d'appui pour la discussion, l'interprétation ou les activités en classe.

Un résultat revient régulièrement dans les commentaires des enseignants. La très forte simplification des VE constitue à la fois leur principal atout et leur principale limite. Les textes sont faciles à comprendre. Certains répondants ont néanmoins le sentiment que cette simplification réduit parfois la profondeur du texte ou atténue ce qui fait son intérêt littéraire. L'accessibilité maximale ne semble donc pas automatiquement conduire au meilleur potentiel pédagogique.

Concernant la place des éléments littéraires dans les processus de simplification, qui constituait un point central dans nos questions de recherche de départ, l'idée était que l'IA pouvait lisser les métaphores, les sous-entendus ou les effets stylistiques pour produire un texte plus correct sur le plan grammatical, mais aussi plus neutre et moins expressif. L'analyse du corpus conduit à un constat différent de celui que nous avons formulé au départ.

Dans les VE, plusieurs procédés littéraires disparaissent plus souvent que prévu. Les métaphores complexes sont régulièrement remplacées par des formulations plus simples. Les effets sonores sont parfois supprimés et certains sous-entendus psychologiques deviennent explicites ou sont condensés. Cette simplification facilite la lecture, mais elle réduit aussi certains effets propres au texte littéraire.

ChatGPT adopte généralement une autre stratégie. Elle ne conserve pas toujours les formulations originales, mais elle cherche souvent à préserver leur fonction. Les images sont reformulées plutôt qu'effacées. Certains indices sensoriels restent

présents. Les dimensions psychologiques ou relationnelles demeurent plus visibles dans plusieurs extraits. Nous n’observons donc pas une suppression systématique de la littérarité par l’IA.

Au fil des analyses, il est devenu de plus en plus difficile de ramener les résultats à une simple opposition entre les productions humaines et celles générées par l’IA. Nous pensions surtout observer deux degrés différents de simplification au départ. Finalement, il apparaît plutôt deux manières distinctes de concevoir ce qu’est un texte adapté. Cette observation ne permet pas encore de trancher sur la meilleure approche. Elle invite surtout à examiner ce que chacune pourrait apporter et ce qu’elle fait perdre.

3.5.2. Forces et limites de la simplification par l’IA

Les résultats des questionnaires mettent en évidence un point qui ne faisait pas partie de nos hypothèses initiales : la lisibilité élevée des textes ne coïncide pas systématiquement avec leur intérêt didactique.

Les VE obtiennent les meilleures notes en matière de clarté. Les indicateurs de lisibilité vont dans le même sens. Les phrases y sont plus courtes, le vocabulaire plus accessible, la compréhension immédiate paraît facilitée. Pourtant, lorsque les enseignants évaluent le potentiel pédagogique ou la capacité du texte à susciter des échanges, ce sont souvent les VIA qui atteignent les scores les plus élevés. Les évaluations recueillies nous ont amenée à envisager la simplification sous un autre angle. Réduire la difficulté d’un passage ne revient pas nécessairement à augmenter tout ce qu’il peut offrir en situation d’apprentissage.

Un document très transparent laisse parfois moins de prise à la discussion. À plusieurs reprises pendant l’analyse, nous avons eu le sentiment que certains passages devenaient plus faciles à comprendre, mais aussi un peu moins stimulants à explorer. Une partie des descriptions disparaît, certains détails sont condensés, plusieurs nuances sont reformulées de manière directe. Le sens devient plus abordé, et la matière à interpréter se réduit aussi. Les VIA occupent une position un peu différente. Elles restent simplifiées, parfois fortement, mais elles conservent davantage d’éléments liés au texte original. Les enseignants semblent avoir été sensibles à cette

présence résiduelle de contenu littéraire lorsqu'ils ont évalué l'intérêt des textes pour la classe.

Au moment d'interpréter ces résultats, un aspect pratique nous a également semblé difficile à ignorer. Adapter un extrait littéraire demande souvent un investissement important de temps. Dans la réalité du terrain, tous les professeurs ne disposent pas des mêmes ressources et n'ont pas toujours accès à des ouvrages déjà adaptés à leur public. La possibilité de produire rapidement plusieurs versions d'un même extrait peut donc représenter un appui non négligeable. Dans cette perspective, l'IA apparaît moins comme un concurrent du travail éditorial que comme un moyen supplémentaire de générer des supports à partir d'un document. La question devient moins celle du remplacement que celle de l'élargissement des possibilités.

Les résultats restent encourageants pour les versions générées par l'IA, mais certaines réponses nous empêchent malgré tout d'adopter une vision trop enthousiaste. Les commentaires ouverts montrent que l'acceptation des VIA n'est pas totale. Certains enseignants mentionnent un style jugé répétitif, des formulations parfois peu naturelles ou un manque de relief émotionnel. Ces remarques restent minoritaires, mais elles nous rappellent que la qualité linguistique d'un texte ne se réduit pas à sa seule correction grammaticale. À la fin de cette analyse, nous avons parfois eu l'impression que la comparaison entre la VE et la VIA était moins une opposition entre réussite et échec qu'une confrontation entre deux manières différentes d'aborder la simplification. Certaines versions privilégient l'accès au sens. D'autres laissent davantage d'éléments à explorer après la première lecture. Les deux logiques ne poursuivent finalement pas exactement le même objectif. Les résultats obtenus ici semblent davantage se situer dans cet espace intermédiaire que dans une hiérarchie claire entre les deux approches.

En relisant l'ensemble des résultats, nous nous sommes posé une question assez simple : quelle part de ces observations appartient réellement au modèle lui-même, et quelle part provient des instructions qui lui ont été fournies ?

3.5.3. Influence de l'ingénierie du prompt sur les résultats obtenus

Pendant le processus d'analyse, une question a fini par prendre plus de place que prévu. Les textes observés nous renseignent-ils réellement sur les capacités de l'IA à simplifier une œuvre littéraire ? Ou décrivent-ils surtout ce qui se produit lorsqu'un modèle reçoit un ensemble très précis d'instructions pédagogiques ?

Les VIA utilisées dans cette recherche ne sont pas apparues spontanément. Avant même la génération des textes, nous avons effectué plusieurs choix : niveau A2, vocabulaire fréquent, maintien de l'intrigue. Le modèle travaillait à l'intérieur de ces limites. Une partie de résultat était donc déjà inscrite dans la demande.

Cela devient particulièrement visible quand nous regardons les indicateurs de lisibilité. La réduction de la longueur des phrases apparaît dans tous les extraits. Au premier regard, il serait tentant d'y voir une caractéristique propre du modèle. Pourtant, c'était précisément ce qui lui était demandé. Même constat pour le lexique : les mots les plus rares ont souvent disparu, mais cette orientation figurait également dans les consignes de départ.

En réalité, plus nous avançons dans les comparaisons, moins la frontière semblait nette entre ce qui venait du modèle et ce qui venait du prompt. Les deux paraissent étroitement liés. Les textes produits ressemblent autant à une réponse de l'IA qu'à la traduction concrète d'une série de choix méthodologique effectués en amont.

Nos hypothèses initiales rendent cette question encore plus intéressante. Nous pensions voir apparaître une banalisation lexicale beaucoup plus marquée. Les résultats racontent autre chose. Des termes relativement spécialisés (« mousseline », « draperie » par exemple) ou certaines descriptions (« une galerie présentait des boutons » dans la VIA4), qui sont jugés difficiles à comprendre pour des apprenants du niveau A2, demeurent présents dans plusieurs extraits. Rien ne permet d'affirmer qu'ils auraient été conservés avec un autre type d'instructions. Un prompt davantage centré sur le vocabulaire aurait sans doute produit un texte différent.

La même impression apparaît lorsque nous observons les éléments littéraires. Plusieurs images, certains effets sensoriels ou des détails liés à l'atmosphère du récit restent visibles dans les VIA alors que nous nous attendions plutôt à les voir s'effacer

progressivement. Le prompt demandait de préserver l'intrigue et les informations importantes. À force de relire les extraits, nous avons fini par nous demander si cette présence partielle de la littérarité n'était pas aussi une conséquence indirecte de cette consigne.

Parler des performances de l'IA comme une propriété stable devient alors délicat. Les résultats présentés ici correspondent à une situation particulière. Un corpus différent. Un autre modèle. Quelques lignes ajoutées ou supprimées dans le prompt. L'équilibre observé entre simplification et conservation du texte source aurait pu évoluer dans une autre direction.

Cette question dépasse finalement notre corpus. Dans la réalité du terrain, rien ne garantit que deux enseignants obtiendront des ressources identiques en utilisant le même outil. Il suffit parfois d'une consigne formulée autrement, d'un objectif pédagogique légèrement différent ou d'une attente plus précise pour que le texte prenne une autre direction.

Nous n'avions pas accordé une place centrale au prompt tout au début du travail. Les analyses nous ont progressivement amenée à revoir cette position. Derrière chaque version générée se cachent des décisions humaines : ce qu'il faut simplifier, ce qu'il faut garder, ce qui paraît essentiel pour un apprenant A2. Le texte final porte aussi la trace de ces arbitrages. À certains moments, il nous a même semblé que comprendre le prompt aidait presque autant à comprendre les résultats que l'étude des textes eux-mêmes.

3.5.4. Vers un modèle de simplification hybride adapté au niveau A2

Après plusieurs allers-retours entre les extraits, il devenait difficile de désigner un véritable gagnant. Certains passages donnaient l'avantage à la VE, puis l'extrait suivant venait nuancer cette impression. À force de comparer les textes, nous avons surtout compris que chaque solution gagnait quelque chose au moment même où elle abandonnait autre chose.

Les VE rassurent en rendant le texte plus transparent et abordable, mais cette simplification est parfois perçue comme excessive, certains enseignants y voyant une perte de relief et une tendance au résumé plutôt qu'au texte littéraire. Les VIA donnent

une impression vraiment distincte. Elles simplifient elles aussi, mais sans aller systématiquement jusqu'à l'effacement, de sorte que subsistent quelques éléments de l'univers zolien ainsi qu'une partie de l'atmosphère du récit. À plusieurs moments, nous avons même eu le sentiment que certains passages conservaient davantage de matière à interpréter que ce que les seuls indices de lisibilité laissaient supposer. Cette présence résiduelle semble expliquer pourquoi plusieurs enseignants y ont vu davantage de possibilités de discussion et d'exploitation en classe.

Une autre question a fini par prendre de la place au fil de l'analyse. Le choix entre VE et VIA n'est peut-être pas la bonne question. Les résultats donnent plutôt l'impression que chacune possède des qualités que l'autre ne maîtrise pas complètement.

Plus les résultats s'accumulaient, moins l'opposition entre VE et VIA paraît suffisante pour rendre compte de ce que nous observions. Une organisation du travail fondée sur la complémentarité paraît alors plus intéressante.

Tableau 17. *Proposition d'un modèle hybride de simplification pour le niveau A2*

Étape	Acteur principal
Sélection de l'extrait littéraire	Enseignant
Première simplification	IA
Vérification linguistique et didactique	Enseignant
Élaboration des activités pédagogiques	Enseignant

L'IA pourrait intervenir en amont du processus. Sa rapidité permet de produire plusieurs propositions de reformulation en peu de temps. L'enseignant ne partirait plus d'une page blanche. Son rôle évoluerait plutôt vers un travail de relecture et d'ajustement. Il lui reviendrait notamment de repérer les passages ambigus, de vérifier si le texte correspond bien au niveau attendu et de décider quels éléments culturels ou stylistiques il est préférable de conserver.

Nous constatons que certains choix sont encore difficiles à confier entièrement à l'IA. Ils ne concernent pas seulement la lisibilité, mais aussi la manière dont le texte sera utilisé en classe. Un enseignant peut préférer conserver un sous-entendu, une

description un peu plus longue ou un terme légèrement plus exigeant s'il estime que ces éléments participent à la compréhension de l'œuvre. Ces choix répondent donc à des objectifs pédagogiques qui vont au-delà de la simple clarté linguistique. Aucune règle ne permet vraiment de trancher ces choix à l'avance. Ils dépendent aussi de ce que l'enseignant souhaite faire du texte une fois en classe. L'idée n'est pas de demander à l'enseignant de corriger à chaque fois l'IA. L'outil pourrait plutôt être envisagé comme une première étape du travail de simplification.

À plusieurs reprises, les commentaires des enseignants nous ont poussée à reconsidérer ce que nous entendions par « simplifier ». Certaines difficultés continuent parfois à avoir une fonction. Elles peuvent nourrir une discussion, attirer l'attention sur un phénomène linguistique ou ouvrir un espace d'interprétation. Au fond, une partie de la difficulté fait parfois encore partie de l'apprentissage. La question n'est alors plus seulement de savoir ce qu'il faut enlever, mais aussi ce qu'il vaut la peine de conserver.

3.5.5. Limites de la recherche

Dans l'ensemble de ce travail, une première limite peut être nommée, c'est que notre analyse repose sur le corpus provenant d'un seul roman littéraire, *Au Bonheur des Dames* d'Émile Zola. Ce choix nous a permis de suivre les transformations avec précision et de suivre les mêmes personnages, les mêmes situations et les mêmes procédés d'écriture à travers les différentes versions simplifiées. Une telle focalisation réduit aussi la portée des résultats. Rien ne garantit que les tendances observées ici se retrouveraient dans un roman d'un autre auteur, dans un texte contemporain ou dans une œuvre où la narration, le style et les enjeux littéraires seraient différents. Ce mémoire permet donc surtout de comprendre comment l'édition pédagogique et l'IA ont simplifié un corpus particulier.

Le protocole expérimental impose lui aussi certaines limites. Les textes artificiels ont tous été générés à partir d'un seul modèle et d'un seul prompt principal. Une partie des choix observés peut provenir du modèle lui-même. Une autre partie peut venir des consignes formulées au moment de la génération. Dans notre cas, le prompt demandait explicitement de viser un niveau A2, d'employer un vocabulaire fréquent

et de conserver l'intrigue principale. Les reformulations produites ne sont donc jamais le simple reflet des capacités de l'outil. Elles résultent d'une rencontre entre un modèle et des instructions précises.

Il devenait de plus en plus difficile d'ignorer le rôle joué par les consignes de génération au fil de l'analyse. Certaines caractéristiques que nous avons d'abord tendance à attribuer à l'IA pourraient en réalité être liées à la manière dont la tâche lui a été présentée. Un prompt davantage centré sur la conservation du style, des images ou des effets littéraires aurait peut-être conduit à des résultats différents. À l'inverse, des consignes plus strictes sur la lisibilité auraient pu rapprocher davantage les VIA des VE.

Un autre élément mérite d'être signalé. Les textes de cette recherche ont été générés avec ChatGPT 4o. Ce choix correspondait à l'outil disponible au moment de la collecte des données. La difficulté est que ce type de modèle évolue en permanence. Une mise à jour peut modifier certaines stratégies de reformulation sans que l'utilisateur en soit nécessairement informé. Deux chercheurs reproduiraient exactement le même protocole quelques mois plus tard pourraient obtenir des textes différents. La situation est un peu inhabituelle. Dans beaucoup de recherches, l'outil reste identique pendant plusieurs années. Ici, l'objet étudié continue d'évoluer pendant que nous écrivons sur lui.

À plusieurs reprises, nous avons également eu le sentiment que certains phénomènes mériteraient d'être observés à une échelle plus large. Les commentaires des enseignants ouvrent de nombreuses pistes qui dépassent le cadre de ce mémoire. Les différences entre locuteurs natifs et non natifs apparaissent par exemple de manière intéressante, mais l'effectif reste trop réduit pour tirer des conclusions solides. Le même constat vaut pour l'influence de l'expérience professionnelle ou du contexte d'enseignement.

Beaucoup de questions restent finalement sans réponse. Une comparaison entre plusieurs modèles d'IA permettrait de déterminer si les tendances observées proviennent réellement de l'IA en général ou d'un système particulier. L'étude de plusieurs prompts, construits selon des logiques différentes, apporterait également

des éléments utiles. La simplification littéraire produite à partir d'une consigne centrée sur la lisibilité n'est probablement pas la même que celle obtenue à partir d'une consigne privilégiant la préservation de la littérarité.

Une absence apparaît assez clairement à la fin de ce travail. Nous avons beaucoup observé les textes et recueilli le regard des enseignants. Les apprenants, eux, restent absents de l'enquête. Or ce sont précisément eux qui lisent les textes simplifiés. Certaines versions jugées intéressantes par les enseignants pourraient ne pas produire les effets attendus auprès d'un public A2 réel. D'autres, au contraire, pourraient être mieux reçues que ce que laissent penser les évaluations professionnelles. Cette question n'a pas trouvé sa place dans le cadre de ce mémoire. Elle constitue probablement l'une des pistes les plus stimulantes pour la suite.

CONCLUSION

Ce travail est parti d'une question qui s'est imposée progressivement au fil de nos lectures et de nos expérimentations. Depuis que les outils d'intelligence artificielle sont devenus accessibles au grand public, ils sont de plus en plus présents dans les pratiques quotidiennes des enseignants. Pourtant, leur place dans l'enseignement de la littérature en FLE reste encore incertaine. Une question revenait constamment au fil de ce travail : une simplification produite par ChatGPT peut-elle réellement constituer une alternative crédible à une adaptation éditoriale élaborée par des spécialistes ? Plus précisément, dans quelle mesure ces deux types de simplification se distinguent-ils lorsqu'il s'agit de rendre *Au Bonheur des Dames* accessible à des apprenants de niveau A2, sans perdre ce qui fait la richesse d'un texte littéraire ?

Pour répondre à cette question, nous avons choisi de mettre face à face deux façons très différentes de simplifier un texte littéraire. L'une est le fruit d'un véritable travail éditorial. Elle a été relue, corrigée, retravaillée, puis pensée pour des apprenants de FLE. L'autre naît en quelques secondes seulement. ChatGPT produit une proposition presque immédiatement. Notre intention n'était pourtant pas de désigner un vainqueur. Ce qui nous intéressait était ailleurs. Nous voulions comprendre ce que chaque démarche conservait du texte de Zola. Et ce qu'elle abandonnait en chemin.

Dans l'ensemble, les indicateurs quantitatifs racontent la même histoire. Dans les deux cas, le texte devient plus facile à lire. Les phrases raccourcissent et le vocabulaire se simplifie. Les indices de lisibilité progressent eux aussi. Mais cette impression d'ensemble ne dure pas longtemps. Dès que l'on entre dans les détails, les différences apparaissent. L'adaptation éditoriale intervient de manière plus radicale : elle coupe davantage. Il est notable que beaucoup de descriptions disparaissent et que le texte devient nettement plus court. À la fin, il ne reste que les éléments indispensables pour suivre l'histoire. ChatGPT suit une autre voie. Le texte qu'il produit reste plus développé, parfois même un peu plus ample qu'on ne l'attendrait dans une adaptation destinée à des apprenants A2. Les descriptions, les personnages secondaires, certains repères de contexte demeurent présents, ce qui oblige parfois à un effort de lecture un peu plus grand, une petite tension supplémentaire, presque

sensible à certains moments, tout en offrant au lecteur davantage d'espace pour entrer dans l'univers du roman et en sentir l'atmosphère.

Les procédés de simplification observés au cours de cette étude confirme cette opposition. Les éditions traditionnelles s'appuient souvent sur la suppression, sur ce geste discret et ferme qui enlève les détails susceptibles de ralentir la compréhension. ChatGPT, lui, intervient autrement. Il reformule davantage qu'il ne supprime. Les mots les plus difficiles sont souvent remplacés, plusieurs phrases sont réécrites et une partie de l'atmosphère du texte est encore préservée. Il en résulte parfois une impression de proximité plus forte avec l'œuvre originale, comme si le texte simplifié n'avait pas complètement quitté son point de départ. Cela crée un équilibre fragile et imparfait, qui se dessine alors entre accessibilité et préservation de l'atmosphère du roman, même si cet équilibre reste parfois instable.

Les résultats de l'expérimentation apportent un éclairage intéressant. Les enseignants interrogés considèrent généralement que la version éditoriale est plus claire pour un public A2. Nous nous attendions à ce résultat. En revanche, la version générée par ChatGPT est plus souvent choisie lorsqu'il est question des activités proposées en classe. Les détails qu'elle garde changent parfois beaucoup de choses. Ils donnent davantage d'idées pour poser des questions, relancer les échanges, parler des personnages ou bien imaginer la suite d'une scène. Finalement, un texte un peu plus riche ne complique pas toujours la lecture, il peut aussi rendre le travail en classe plus dynamique.

Ces observations nous ramènent aussi à quelque chose que nous connaissons bien, l'enseignement du FLE au Vietnam. Dans beaucoup de classes, le temps manque. On avance vite, parfois trop vite. Il faut finir le programme, préparer les examens, garder les apprenants motivés. Dans ces conditions, la littérature passe souvent au second plan, non pas parce qu'elle n'a pas de valeur, mais parce qu'elle paraît difficile à faire entrer dans le rythme réel du cours. Cette réalité explique sans doute pourquoi les textes littéraires restent parfois peu exploités dans les cours de FLE. Nous la voyons même comme assez normale. Les outils d'intelligence artificielle peuvent alors aider, au moins pour ouvrir une première porte. Ils permettent de proposer rapidement une

version simplifiée, puis de la reprendre, de la corriger, de l'adapter à ce que l'enseignant veut vraiment faire en classe. Mais nous restons convaincue que cette rapidité ne suffit pas. Elle ne remplace pas le regard de l'enseignant, ni son intuition, ni sa manière de sentir ce qui peut fonctionner avec un groupe précis. Un texte simplifié n'est pas seulement un texte plus facile à lire. C'est aussi une certaine façon d'entrer dans la littérature, de la rendre proche sans la vider de sa substance.

Ce travail comporte naturellement plusieurs limites. Le corpus repose sur une seule œuvre d'Émile Zola. Les résultats auraient sans doute été différents avec un autre auteur, une autre époque ou un autre genre littéraire. Les textes générés par l'intelligence artificielle proviennent d'un seul modèle et d'un seul prompt. Par ailleurs, les performances des modèles d'intelligence artificielle évoluent très rapidement. Les observations formulées dans ce mémoire correspondent donc à un état des technologies au moment de la réalisation de cette recherche et gagneraient à être réévaluées dans quelques années. L'expérimentation repose enfin sur le regard d'enseignants de FLE. Les apprenants, qui sont pourtant les premiers concernés, n'ont pas participé directement à cette évaluation. Leur expérience de lecture pourrait parfois confirmer ces résultats, parfois les nuancer.

Cette recherche ouvre finalement plus de questions qu'elle n'en résout. Plusieurs pistes mériteraient d'être explorées. Nous pensons notamment à une comparaison entre différents modèles d'intelligence artificielle, dont les comportements peuvent varier sensiblement. Le rôle du prompt constitue également un point important : au cours de cette étude, nous avons d'ailleurs constaté que quelques modifications dans les consignes pouvaient entraîner des différences parfois notables entre les textes produits. Il serait également intéressant d'observer la réception de ces versions simplifiées directement auprès d'apprenants de niveau A2, afin de mieux comprendre leurs difficultés de lecture, mais aussi leur motivation et leur plaisir à lire. Enfin, cette réflexion pourrait être prolongée avec d'autres niveaux du CECRL ou à partir d'œuvres appartenant à des genres littéraires différents.

Si cette recherche nous conduit finalement à une conclusion, ce n'est peut-être pas seulement sur l'intelligence artificielle, mais sur la simplification littéraire elle-même.

Au fil de ce travail, nous avons progressivement compris qu'une simplification ne commence pas avec les mots. Elle commence avec les choix que l'on accepte de faire. Simplifier ne revient pas seulement à rendre un texte plus facile. Cela oblige à choisir : garder certains éléments, en laisser d'autres de côté. Rien n'est complètement neutre. Chaque choix modifie un peu la manière dont le lecteur entre dans l'œuvre. L'enseignant essaie donc de trouver un équilibre qui reste souvent fragile : aider les apprenants à comprendre sans effacer ce qui fait encore entendre la voix du texte. Quant à l'intelligence artificielle, elle peut accompagner cette démarche plutôt que la remplacer. Elle peut proposer une première version, ouvrir des pistes, faire gagner du temps, mais elle ne dispense jamais de ce travail de réflexion et d'ajustement qui accompagne toute réécriture destinée à la classe. Au fond, ce mémoire nous rappelle que la valeur d'un texte simplifié ne tient pas seulement à son origine, qu'elle soit humaine ou générée par l'IA, mais à la qualité des choix qui ont guidé sa transformation. C'est peut-être là que réside le véritable enjeu de la simplification en FLE : non pas produire un texte plus simple à tout prix, mais permettre aux apprenants d'entrer dans une œuvre littéraire sans rompre le dialogue qu'elle entretient avec ses lecteurs.

BIBLIOGRAPHIE

- Barthes, R. (1973). *Le plaisir du texte*. Seuil.
- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency*, 610-623.
- Bhatia, V. K. (1983). Simplification v. easification—The case of legal texts¹. *Applied linguistics*, 4(1), 42-54.
- Blum-Kulka, S. (1986). Shifts of cohesion and coherence in translation. *Interlingual and intercultural communication: Discourse and cognition in translation and second language acquisition studies*, 17, 17-35.
- Blum, S., & Levenston, E. A. (1978). Universals of lexical simplification. *Language learning*, 28(2), 399-415.
- Brouwers, L., Bernhard, D., Ligozat, A.-L., & François, T. (2012). Simplification syntaxique de phrases pour le français (Syntactic Simplification for French Sentences)[in French]. Proceedings of the Joint Conference JEP-TALN-RECITAL 2012, volume 2: TALN, Grenoble.
- Bruner, J. (1983). *Le développement de l'enfant: savoir faire, savoir dire*. Presses Universitaires de France.
- Brunet, É. (1985). Les Rougon-Macquart. Aspects quantitatifs. *Revue Informatique et Statistique dans les Sciences Humaines*, XXI(1-4), 35-52.
<https://hal.science/hal-01465289>
- Brunet, É. (2016). La phrase de Zola. *Texto! Textes et Cultures*, 21(1), Publication électronique. <https://shs.hal.science/halshs-01367627/>
- Buet, F. (2022). *Modèles neuronaux pour la simplification de parole, application au sous-titrage Paris-Saclay*. Paris. <https://theses.hal.science/tel-03920729/>
- Cardon, R. (2021). *Simplification automatique de textes techniques et spécialisés* Université de Lille].

- Causa, M. (2001). De la simplification en classe de français, langue professionnelle. *Les Carnets du Cediscor. Publication du Centre de recherches sur la didacticité des discours ordinaires*(7), 65-79.
- Charaudeau, P. (2009). *L'identité culturelle entre soi et l'autre*. Actes du colloque de Louvain-la-Neuve (2005), Louvain-la-Neuve. <https://www.patrick-charaudeau.com/L-identite-culturelle-entre-soi-et-l-autre.html>
- Clamens, P. A. (1959). Style Descriptif chez Loti et Chez Zola-Perspective Statique et Perspective Cinétique. *Modern Language Notes*, Vol. 74, No. 6, 521-530.
- Claustres, F. (2023). *Au bonheur des dames Lectures CLE en Français Facile*. CLE International.
- Collès, L. (2002). Les textes en français simplifié: une voie pour l'enseignement de la littérature. *Praxis*, 4(1), 402-409.
- Conseil de l'Europe. (2001). Cadre européen commun de référence pour les langues. Paris, France, Didier.
- Cuq, J.-P., & Gruca, I. (2005). *Cours de didactique du français langue étrangère et seconde* (4^e ed.). Presses universitaires de Grenoble.
- Cuq, J. P. (2003). *Dictionnaire de didactique du français langue étrangère et seconde*. CLE International. <https://books.google.be/books?id=XMpwQgAACAAJ>
- Daelemans, W., Höthker, A., & Sang, E. F. T. K. (2004). Automatic Sentence Simplification for Subtitling in Dutch and English. LREC,
- David, J. (2010). La didactique de la littérature: impasses et promesses. *Cahiers du Centre de Linguistique et des Sciences du Langage*(27), 77-90.
- Defays, J.-M., & Deltour, S. (2003). *Le français langue étrangère et seconde: enseignement et apprentissage* (Vol. 251). Mardaga.
- Delbart, A.-R. (1994). Textes en français simplifié: simplicité ou simplisme? *Enjeux (Namur)*, 31, 133-147.
- Dubois, J. (1968). Zola chez un public d'aujourd'hui. *Europe*, Vol. 46(468), 257-266.
- Dufays, J.-L. (2017). Apprendre la littérature, pour quoi faire ? Du débat des experts aux discours des élèves. *Transpositio*, 1, 1-14.

- Dufays, J.-L. (2020). Comment et pourquoi développer la compétence de lecture littéraire ? *Echanges - Bulletin de formation continuée des professeurs de français*, no.39, 4-8. <http://hdl.handle.net/2078.1/227869>
- Eco, U. (1965). *L'œuvre ouverte*. Seuil. <https://ils.bib.uclouvain.be/global/documents/1697922>
- Flesch, R. (1948). A new readability yardstick. *Journal of applied psychology*, 32(3), 221.
- Fuchs, C. (2020a). Paraphrase et reformulation : un chassé-croisé entre deux notions. *Autour de la reformulation*, 36, 41-55. <https://hal.science/hal-03044034> (Coll. Recherches et Rencontres)
- Fuchs, C. (2020b). Paraphrase et reformulation: un chassé-croisé entre deux notions. *Autour de la reformulation*, 36, 41-55.
- Gagnon, J. C. (1983). *Lire une bande dessinée: pratique, théorie, pratiques*. Editions Ville-Marie.
- Gala, N., François, T., Javourey-Drevet, L., & Ziegler, J. C. (2018). La simplification de textes, une aide à l'apprentissage de la lecture. *Langue française*, 199(3), 123-131.
- Gala, N., Todirascu, A., Bernhard, D., Wilkens, R., & Meyer, J.-P. (2020). Transformations syntaxiques pour une aide à l'apprentissage de la lecture: typologie, adéquation et corpus adaptés. *SHS Web of Conferences*, 78, 14006.
- Ganascia, J.-G. (2022). Intelligence artificielle et épistémologie. Allers-retours indispensables. JECIS,
- Gerolimich, S., & Vecchiato, S. (2014). Pas simple de simplifier ! Pratiques de « réécriture » des documents médicaux orientés vers l'utilisateur. In *Repenser le rôle des pratiques langagières dans la constitution des espaces sociaux contemporains* (pp. 35-65). Louvain-la-Neuve : De Boeck Supérieur.
- Giasson, J. (1996). *La compréhension en lecture*. De Boeck Université.
- Hamon, P. (1993). *Du descriptif*. Paris: Hachette supérieur.
- Hijazi, R. (2020). Transformations syntaxiques entre niveaux de simplification dans le corpus Newsela (Syntactic transformations between simplification levels in

- the Newsela corpus). Actes de la 6e conférence conjointe Journées d'Études sur la Parole (JEP, 33e édition), Traitement Automatique des Langues Naturelles (TALN, 27e édition), Rencontre des Étudiants Chercheurs en Informatique pour le Traitement Automatique des Langues (RÉCITAL, 22e édition). Volume 3: Rencontre des Étudiants Chercheurs en Informatique pour le TAL., Nancy, France.
- Hoàng, V. V. (2007). Vai trò của giáo viên và học sinh ngoại ngữ ở trường học phổ thông trong lớp học theo đường hướng lấy người học làm trung tâm. *VNU Journal of Foreign Studies*, 23(2), 120.
- Iser, W. (1985). *L'acte de lecture: théorie de l'effet esthétique* (E. Sznycer, Ed.). Mardaga. <https://ils.bib.uclouvain.be/global/documents/194214>
- Jauss, H. R. (1990). *Pour une esthétique de la réception* (C. Maillard & J. Starobinski, Eds.). Gallimard. <https://ils.bib.uclouvain.be/global/documents/278772>
- Jenny, L. (2010). N'imaginez jamais : pour une pédagogie littéraire de l'immersion. *Les 11 es Rencontres des chercheurs en didactique des littératures*.
- Ji, W., Yuan, W., Getzen, E., Cho, K., Jordan, M. I., Mei, S., Weston, J., Su, W. J., Xu, J., & Zhang, L. (2026). An overview of large language models for statisticians. *The American Statistician*(just-accepted), 1-106.
- Julia, L. (2019). *L'intelligence artificielle n'existe pas*. First.
- Kandel, L., & Moles, A. (1958). Application de l'indice de Flesch à la langue française. *Cahiers Etudes de Radio-Télévision*, 19(1958), 253-274.
- Karsenti, T. (2018). Intelligence artificielle en éducation: L'urgence de préparer les futurs enseignants aujourd'hui pour l'école de demain. *Formation et profession*, 26(3), 112-119. <https://doi.org/10.10.18162/fp.2018.a159>
- Kayembe, E. (2019). L'exploitation du texte littéraire en classes débutantes de FLE. AATF convention,
- Kintsch, W. (1982). Aspects de la compréhension de textes. *Bulletin de psychologie*, 35(356), 777-787. <https://doi.org/10.3406/bupsy.1982.12021>
- Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *Relc Journal*, 54(2), 537-550.

- Konti, K., & Varsamidou, A. (2025). Perceptions des enseignants de FLE sur l'intégration de ChatGPT et d'autres outils d'intelligence artificielle dans l'enseignement du FLE : enjeux et perspectives. *European Journal of Foreign Language Teaching*, 9(3).
- Krashen, S. (2013). The case for non-targeted, comprehensible input. *Journal of Bilingual Education Research & Instruction*, 15(1), 102-110.
- Laidoudi, A., & Messaour, R. (2024). Adaptation linguistique et culturelle des textes littéraires en classe de FLE : Quels effets sur la compréhension ? *Akofena*, 1(11), 355-368.
- Langlade, G. (2014). La lecture subjective est-elle soluble dans l'enseignement de la littérature? *Études de lettres*(1), 47-64.
- Laurent, J.-M. (2008). Être un enseignant médiateur? *La nouvelle revue de l'adaptation et de la scolarisation*, 42(2), 15-32.
- Lebrun, M. (1993). L'école et le plaisir de lire. *Documentation et bibliothèques*, 39(3), 111-116. <https://www.erudit.org/en/journals/documentation/1993-v39-n3-documentation01710/1028746ar/abstract/>
- Leroy, G., Kauchak, D., Harber, P., Pal, A., & Shukla, A. (2024). Text and audio simplification: Human vs. chatgpt. *AMIA Summits on Translational Science Proceedings*, 2024, 295.
- Likert, R. (1932). A technique for the measurement of attitudes. *Archives of psychology*, 22, 5-55.
- Lovrečki, K. (2023). *Textes littéraires adaptés en classe de FLE: analyse du roman « Vingt mille lieues sous les mers » de Jules Verne* Université de Zagreb].
- Maingueneau, D. (2004). *Le discours littéraire*. Paris: Armand Colin.
- Mansouri, K. (2025). La Syntaxe chez Émile Zola: une écriture des forces, du rythme et du réel. *Contextes Didactiques, Linguistiques et Culturels*, 3(3), 139-157.
- Martinot, C. (2010). Reformulation et acquisition de la complexité linguistique. *Travaux de linguistique*, 61(2), 63-96.
- Matthews, J. H. (1957). *Les deux Zola: Science et personnalité dans l'expression*. Geneva: Droz.

- Melançon, J. (1997). Préface. In M. Noël-Gaudreault (Ed.), *Didactique de la littérature: bilan et perspectives* (pp. 5-11). Nuit Blanche.
- Mitterand, H. (1985). Le naturalisme théorique de Zola in *Recherches à Valenciennes-Journées Zola-Mai 1984. Lez Valenciennes*(10), 181-193.
- Mitterand, H. (1990). *Zola: l'histoire et la fiction*. Presses universitaires de France (PUF).
- Morin, V., & Raynault, A. (2023). La culture informationnelle et l'arrivée de l'Open AI. *Forumlecture. ch*, 3. <https://www.forumlecture.ch/lffl/2023/3/807>
- Nisioi, S., Štajner, S., Ponzetto, S. P., & Dinu, L. P. (2017). Exploring neural text simplification models. *Proceedings of the 55th annual meeting of the association for computational linguistics (volume 2: Short papers)*,
- Numa-Bocage, L. (2007). La médiation didactique: un concept pour penser les registres d'aide de l'enseignant. *Carrefours de l'éducation*, 23(1), 55-70.
- Paillé, P., & Mucchielli, A. (2021). *L'analyse qualitative en sciences humaines et sociales-5e éd.* Armand Colin.
- Pennac, D. (1992). *Comme un roman*. Gallimard.
- Puren, C. (2006). De l'approche communicative à la perspective actionnelle. *Le français dans le monde*, 347, 37-40.
- Rinck, F. (2025). Des textes de l'IA générative pour former à l'enseignement de la production écrite. *Revue internationale des technologies en pédagogie universitaire (RITPU)*, 22(1). <https://doi.org/10.18162/ritpu-2025-v22n1-10>
- Ronveaux, C., & Schneuwly, B. (2018). *Lire des textes réputés littéraires: disciplination et sédimentation*. PIE Peter Lang.
- Rouxel, A. (2007). Pratiques de lecture: quelles voies pour favoriser l'expression du sujet lecteur? *Le français aujourd'hui*, 157(2), 65-73.
- Safira, J. A., Widayanti, D. V., & Handayani, S. (2019). La simplification des fables françaises pour les lycéens. *Didacticofrancia: Journal Didactique du FLE*, 8(1), 56-61.
- Saint-Gérard, J.-P. (1993). Philippe Hamon, La Description littéraire. De l'Antiquité à Roland Barthes: une anthologie. *Romantisme*, 23(79), 116-117.

- Schreiber, M. (2009). Les procédés de l'adaptation diminutive: Balzac, Flaubert, Verne et Zola simplifiés, résumés, purifiés. *Cahiers d'Études Germaniques*, 56(1), 91-99.
- Sweller, J., Chanquoy, L., & Tricot, A. (2007). *La charge cognitive: Théorie et applications*. Armand Colin.
- Tôn, Q. C. (2025). PHÁT TRIỂN KHUNG NĂNG LỰC SỬ DỤNG TRÍ TUỆ NHÂN TẠO TRONG GIÁO DỤC. *Tạp chí Tâm lý - Giáo dục*, 31(04), 25-28.
- UNESCO. (2024). *Orientations pour l'intelligence artificielle générative dans l'éducation et la recherche*. UNESCO Publishing.
- Vaishya, R., Misra, A., & Vaish, A. (2023). ChatGPT: Is this version good for healthcare and research? *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, 17(4), 102744. <https://doi.org/10.1016/j.dsx.2023.102744>
- Vandendorpe, C. (1989). *Apprendre à lire des fables*. Montréal: Éditions du Préambule.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* (Vol. 86). Harvard university press.
- Widdowson, H. G. (1978). *Teaching language as communication*. Oxford university press.
- Wilkinson, C., Oppert, M., & Owen, M. (2024). Investigating academics' attitudes towards ChatGPT: A qualitative study. *Australasian Journal of Educational Technology*, 40(4), 105-120. <https://doi.org/10.14742/ajet.9456>
- Yang, S. (2025). GPT-4o-based Multimodal Artificial Intelligence Framework to Enhance Personalized English Language Learning. 2025 3rd International Conference on Data Science and Information System (ICDSIS),
- Zhu, Z., Bernhard, D., & Gurevych, I. (2010). A monolingual tree-based translation model for sentence simplification. Proceedings of the 23rd International Conference on Computational Linguistics (Coling 2010),
- Zola, É. (1868/2012). Préface de la Deuxième édition. In *Thérèse Raquin*. Le livre de poche.
- Zola, É. (1873). *La fortune des Rougon*. Charpentier.

Zola, É. (1880). *Le roman expérimental*. Charpentier.

Zola, É. (1883/2024). *Au bonheur des dames*. Folio.

UNIVERSITÉ CATHOLIQUE DE LOUVAIN

Faculté de philosophie, arts et lettres

Place Cardinal Mercier, 31, 1348 Louvain-la-Neuve, Belgique | www.uclouvain.be/fial