A critical study of ChatGPT's potential to support (novice) language teachers in the Wallonia-Brussels Federation

Auteur : Kokkinis, Konstantina
Promoteur(s) : Simons, Germain; Brems, Lieselotte
Faculté : Faculté de Philosophie et Lettres
Diplôme : Master en langues et lettres modernes, orientation générale, à finalité didactique
Année académique : 2024-2025
URI/URL : http://hdl.handle.net/2268.2/23094

Université de Liège
Faculté de Philosophie et Lettres
Département de Langues modernes : littérature, linguistique, traduction

A CRITICAL STUDY OF CHATGPT'S POTENTIAL TO SUPPORT (NOVICE) LANGUAGE TEACHERS IN THE WALLONIA-BRUSSELS FEDERATION

Mémoire présenté par Konstantina KOKKINIS
en vue de l'obtention du grade de
Master en Langues et lettres modernes,
Orientation générale à finalité didactique

Promoteur : (Prof.) Germain SIMONS

Année académique 2024/2025

Critères de qualité des travaux de fin d'études de la filière en Langues et lettres modernes

---

1. Questions/thématiques de recherche

- La question de recherche est-elle clairement définie ?
- La question de recherche est-elle originale et/ou scientifiquement ambitieuse ?
- Dans quelle mesure contribue-t-elle à la littérature scientifique et à l'état des connaissances de la discipline ?

2. Mobilisation de la théorie

- Utilisation de sources pertinentes ?
  - Le travail contient-il des références solides et pertinentes ?
  - Le travail contient-il un nombre suffisant de références scientifiques ?
    - Le seuil minimum est fixé à *10 références scientifiques* (à savoir : ouvrage, monographie, article de revue scientifique, chapitre d'ouvrage, compte-rendu…) ; ne comptent pas comme références scientifiques : les articles de blogs et les pages issues de sites de vulgarisation.
- Utilisation pertinente et critique des sources ?
  - Les sources sont-elles mobilisées de manière adéquate dans le texte ?
  - Les citations sont-elles mobilisées de manière pertinente dans le texte ?
  - Les différentes sources sont-elles mises en relation ?
- Les concepts pertinents pour la question de recherche sont-ils clairement définis et maîtrisés ?
- La/Les questions de recherche (et les hypothèses éventuelles qui en découlent) sont-elles pertinentes, principalement en lien avec l'état de l'art ?

3. Méthodologie

- La méthodologie déployée permet-elle de répondre aux questions de recherche ?
- La méthodologie déployée est-elle décrite avec clarté et de manière complète ?
- Le cas échéant : la collecte des données (corpus, échantillon, questionnaire, sources textuelles…) a-t-elle été effectuée de manière rigoureuse ?
- Permet-elle d'apporter des éléments de réponse aux questions de recherche et aux objectifs du travail, et, le cas échéant, de confirmer ou d'infirmer les hypothèses de travail ?

4. Analyse/Commentaire/Résultats

- La présentation des résultats ou observations se base-t-elle sur des preuves textuelles, des citations, des analyses de corpus, des extraits d'entretiens… ?
- Le corpus de travail est-il analysé de manière complète et systématique ?
- Le cas échéant : la base de données a-t-elle été constituée avec rigueur et précision ?
- Les résultats sont-ils présentés de manière claire et précise ?
- Les résultats sont-ils présentés de manière logique, de façon à développer un raisonnement cohérent ?
- Les résultats permettent-ils de répondre aux questions de recherche et de vérifier les hypothèses de travail ?
- Le commentaire permet-il une analyse en lien avec le cadre théorique défini ?

5. Discussion, synthèse, perspectives

- Les observations principales du travail sont-elles résumées de manière claire et mises en relation avec la littérature scientifique ?
- Des pistes de développement sur la base des conclusions principales (pour des recherches futures) sont-elles proposées ?
- Un regard critique sur la démarche mise en œuvre dans le travail est-il proposé ?

6. Qualité de la langue

Il est attendu que le TFE soit rédigé en langue étrangère et que la qualité de la langue mobilisée soit conforme aux attentes académiques. Indépendamment du contenu, le jury a la possibilité de remettre en cause la réussite du travail s'il estime que la qualité de la langue est insuffisante.

- La langue utilisée dans le travail respecte-t-elle les normes orthographiques, grammaticales et syntaxiques ?
- La terminologie scientifique est-elle mobilisée de manière appropriée ?
- Le texte est-il structuré de manière cohérente ?
- Le document respecte-t-il les caractéristiques du style académique ?
- La qualité de rédaction est-elle de nature à remettre en cause la réussite du travail ?

7. Mise en page et typographie

- La présentation matérielle du mémoire (structure, mise en page, typographie) est-elle soignée ?

- La longueur du travail est-elle conforme aux consignes ?

8. Référencement bibliographique et citations

- Toutes les références traitées dans le texte sont-elles présentes dans la bibliographie ?
- Toutes les références présentes dans la bibliographie sont-elles traitées dans le texte ?
- Les normes de citation sont-elles respectées ?
- Les normes bibliographiques sont-elles appliquées de manière cohérente et systématique ?
- Le travail ne contient-il pas de plagiat ; tout propos ne relevant pas d'une réflexion personnelle de l'étudiant·e est-il référencé ?

9. Défense orale

La défense orale permet au jury de vérifier la maîtrise des sujets abordés dans le travail ainsi que l'appareil méthodologique déployé. Elle permet de vérifier les compétences de présentation des étudiant·es et leur aptitude à répondre à des remarques critiques. La défense est publique et se déroule dans la langue étrangère.

Lors de la défense orale, l'étudiant·e propose une synthèse du travail soulignant les résultats principaux, approfondit un aspect particulier de celui-ci ou exploite une thématique connexe. Cette présentation dure au maximum 10 minutes.

- Le contenu de l'exposé est-il présenté de manière concise ?
- L'exposé est-il présenté de manière cohérente ?
- L'étudiant·e répond-il/elle aux critiques et questions de manière adéquate et convaincante ?
- La maîtrise de la langue orale est-elle conforme aux exigences académiques ?
- La langue mobilisée lors de la défense respecte-t-elle les normes grammaticales et lexicales ?

10. Déclaration d'authenticité relative à l'utilisation de l'intelligence artificielle générative

- L'utilisation de plateformes d'intelligence artificielle générative est-elle conforme à ce qui est indiqué dans la déclaration d'authenticité ?

11. Longueur

*La longueur attendue pour un **TFE du master 120** (avec une fourchette de 10 % vers le haut ou vers le bas) est de **240 000 caractères espaces compris**, hors bibliographie et annexes. À titre indicatif, cela correspond à 36 000 mots, hors bibliographie et annexes.*

*La longueur attendue pour un **TFE du master 60** (avec une fourchette de 10 % vers le haut ou vers le bas) est de **160 000 caractères espaces compris**, hors bibliographie et annexes. À titre indicatif, cela correspond à 24 000 mots, hors bibliographie et annexes.*

- La longueur du TFE est-elle conforme aux dispositions réglementaires ?

LIÈGE université
Philosophie et Lettres
LLMT

Déclaration d'authenticité

Je, soussigné·e Konstantina KOKKINIS déclare avoir rédigé le présent travail de fin d'études de manière autonome, sans l'aide non autorisée de tiers et ne pas avoir utilisé d'autres moyens que ceux indiqués. J'ai mentionné, en précisant la source, les passages de ce travail empruntés textuellement ou sous forme de paraphrase à d'autres ouvrages.

Je déclare avoir pris connaissance de la charte ULiège d'utilisation des intelligences artificielles génératives dans les travaux universitaires (https://www.student.uliege.be/cms/c_19230399/fr/faq-student-charte-uliege-d-utilisation-des-intelligences-artificielles-generatives-dans-les-travaux-universitaires) et des restrictions propres à ma filière d'étude, et je déclare que mon travail implique (cochez la case appropriée) :

- ☐ Aucun usage de l'IA générative
- ☑ Un usage de l'IA générative comme assistant linguistique (amélioration de la formulation, de la mise en forme de textes que j'ai rédigés ; cette utilisation est comparable aux correcteurs d'orthographe et de grammaire existants).
- ☐ Un usage de l'IA générative comme assistant à la recherche d'information (aide comparable à l'usage des moteurs de recherche existants qui facilitent l'accès à la connaissance d'un sujet).

Ce travail peut être vérifié pour le plagiat et l'utilisation des intelligences artificielles génératives à l'aide du logiciel approprié. Je comprends qu'une conduite contraire à l'éthique peut entraîner une sanction.

Lieu, date Signature

Liège, le 21/05/2025 

Acknowledgments

First of all, I would like to express my most sincere gratitude to my supervisor, Professor Germain Simons, for his constant support, availability, and valuable guidance throughout this dissertation and during my traineeships. His encouragement, feedback, and presence at every stage made a real difference in both my academic progress and personal development. I feel fortunate to have been mentored by someone so dedicated and supportive.

I also warmly thank Professor Lieselotte Brems, my co-supervisor, for her time, availability and feedback throughout this process.

A special thank you goes to Audrey Renson for her helpful advice on how to design the survey, and to the assistants in the Department of Modern Languages at the University of Liège (Julie Vanhoof, Florence Van Hoof, Lisa Svennberg, and Sophie Radelet) for their useful feedback, which helped me improve the questionnaire.

I am also very grateful to my friend, Lisa Gayet, with whom I co-developed the first part of the survey. More than just a collaborator, Lisa has been by my side since the beginning of university. Her support and energy made this experience not only productive but also enjoyable.

I would like to express my deepest gratitude to my partner, for being by my side every step of the way. His love, patience, and constant belief in me gave me strength when I needed it most.

Finally, I want to thank my family and friends, whose support and presence helped me stay motivated during this journey.

Table of contents

List of Screenshots, Tables and Figures

Screenshots:

Tables:

Figures:

List of appendices

Appendix 1: Evaluation of a didactic sequence produced by the free version of ChatGPT

Appendix 2: Evaluation of a didactic sequence produced by the premium version of ChatGPT

Appendix 3: ChatGPT's didactic sequence – free version with guidance

Appendix 4: ChatGPT's didactic sequence – premium version with guidance

Appendix 5: ChatGPT's didactic sequence – free version without guidance and without memory

Appendix 6: ChatGPT's didactic sequence – free version without guidance, with memory

Appendix 7: A survey of (future) language teachers in Liege: uses and perceptions of ChatGPT

Appendix 8: Usage perspectives: stereotypes in ChatGPT

Appendix 9: Usage perspectives: Asking ChatGPT to generate exercises

Appendix 10: Usage perspectives: ChatGPT's lack of emotional sensitivity

Introduction

As a young language teacher and researcher, I quickly realised that one of the biggest challenges in education today is not knowing *what* to teach, but rather *how* to teach it. In the French-speaking Community of Belgium, several official resources guide our work in language education, such as the CEFR (*Common European Framework of Reference for Languages*), the UAAs (Unités d'Acquis d'Apprentissage), the learning strategies, and the Wallonia-Brussels Federation (WBF) triplet (*connaître, appliquer, transférer*). These documents define learning goals and the skills students should develop. However, they do not offer enough practical steps or methods to help teachers plan and structure their sequences in a clear and consistent way.

The WBF triplet is the only element that hints at a teaching progression. Yet, in practice, this triplet is minimally explained and rarely supported by illustrative examples. Moroever, both the triplet and the UAAs are designed to apply across all educational departments, not specifically for language education. As a result, many young language teachers find themselves with theoretical goals but no structured tools to help translate them into coherent didactic sequences.

In this context, the course "Didactics of Modern Languages" at the University of Liège plays an important role. In this course, among other tools and approaches, we were introduced to the PFE model—Presentation, Fixation, Exploitation— (Simons, 2023) which gives a clear structure for building language sequences. This model is directly linked to the PPP approach (Present, Practice, Produce) used in communicative language teaching, and it fits well with the WBF triplet (Simons, 2023: 30). The PFE model offers future teachers a helpful and concrete way to design their didactic sequences, namely, a series of lessons structured around various learning activities that gradually lead to a final communicative task (Simons, 2023: 28). In comparison, a model refers here to a teaching scenario composed of successive teaching-learning phases, providing a general template that can be adapted to different pedagogical contexts (2023: 28). Personally, the PFE model has become a key tool in my own approach to teaching.

Even with this structure, however, planning lessons is time-consuming, especially for young teachers who are still learning how to manage their time and prepare quality materials. In recent years, new digital tools such as ChatGPT have become more popular, and many teachers—myself included—have started wondering whether this kind of artificial intelligence could be used to help with didactic sequence design. Could ChatGPT reduce the workload? Could it help

generate ideas or save time when creating materials? And more importantly, can it support us without taking away our role as teachers?

These questions are at the heart of this dissertation. My main objective is to find out whether ChatGPT can help young language teachers, like myself, in preparing effective didactic sequences. To do so, I analyse both the tool's ability to create coherent sequences and the perceptions of (future) teachers who use or have tested it. I also explore how ChatGPT fits within the PFE model and whether its outputs can match the requirements of the curriculum.
This work is both a personal and professional investigation. It reflects my interest in using new technologies in education, but also my desire to understand the risks, limitations, and best practices that come with them. Through this study, I hope to provide useful insights for teachers, trainers, and institutions who are asking the same question: *can ChatGPT be a useful ally for language teachers in today's classrooms?*

To guide this investigation, I formulated six main research questions:

1. **What are the (perceived) pedagogical benefits and limitations of using ChatGPT in language education?**
2. **What do (future) language teachers in Liège say about their use of ChatGPT?**
3. **How do (future) language teachers perceive and evaluate the quality of ChatGPT-generated content?**
4. **What training do (future) language teachers receive or expect regarding ChatGPT?**
5. **How do personal and contextual factors influence attitudes and usage of ChatGPT?**
6. **Is ChatGPT capable of creating a coherent didactic sequence, adapted to language education in the Wallonia-Brussels Federation, when guided by a language teacher?**

Each of these questions is accompanied by several sub-questions, which will be introduced at the beginning of the relevant chapters. These sub-questions are there to guide the analysis step by step and offer a clearer, more focused look at how ChatGPT is being used in language education.

The structure of this dissertation reflects this dual focus: on the one hand, a critical evaluation of ChatGPT's pedagogical output; on the other hand, an exploration of how (future) language teachers perceive and engage with the tool. First, the paper opens with a literature review, which explores the historical evolution of AI, its applications in education, and the specific capabilities and challenges of ChatGPT in language education. This section also includes pedagogical recommendations and a reflection on the importance of critical thinking and equity in AI use.

Next, this work examines the official documents that structure language education in the French-speaking Community of Belgium, including the CEFR, UAAs and frameworks. While these resources define learning objectives and competences, they often lack detailed guidance. To address this, the PFE model is introduced as a complementary tool to support the design of coherent didactic sequences and to explore how ChatGPT might assist in this process.

In the third part, I evaluate the didactic sequences produced by ChatGPT under various conditions (free version, premium version, with guidance, without guidance, with and without memory options), using evaluation grids specifically created for this analysis, based on the PFE model. This analysis allows me to assess ChatGPT's pedagogical coherence, its strengths and weaknesses in real teaching scenarios.

The fourth section presents the results of a survey conducted among (future) language teachers in Liège. The questionnaire explores their usage habits, attitudes, and opinions on ChatGPT in education.

Finally, this dissertation concludes with a discussion, followed by a synthesis of key findings, and proposes practical and research perspectives for future integration of AI in language education.

1. Literature review

Before starting the literature review, it is essential to clarify the scope and aims of this section. The purpose of the literature review is to provide a synthesis of existing research related to Artificial Intelligence (AI) and its applications, particularly in the context of language education. This part will give an overview of the current state of research, tracing the development and historical evolution of AI, examining its integration into the field of education, and offering a detailed explanation of ChatGPT. Moreover, the literature review will evaluate the benefits and limitations of ChatGPT, focusing on what it can offer and the difficulties it can cause in language learning. Several key themes will be addressed, including the potential advantages of incorporating ChatGPT into language teaching, as well as the risks and challenges that may arise from its use. In addition, pedagogical recommendations will be proposed, with particular attention given to strategies for fostering critical thinking and ensuring equity in educational practices. By presenting a balanced analysis of both the opportunities and constraints associated with ChatGPT, this literature review seeks to provide a nuanced and critical understanding of what AI means for education today. It is also important to consider that the views presented in academic literature are not entirely neutral. Authors' institutional roles, research interests, and technological orientations may influence their evaluations of ChatGPT, potentially leading to optimism, caution, or skepticism depending on their background and professional experience. Finally, some sections of this literature review will directly contribute to answering specific research questions (and sub-questions) formulated in this paper. Each relevant sub-question will be introduced at the beginning of the corresponding section in a box to guide the reader through the analysis. Here are all the questions that will be answered:

RQ1: ***What are the (perceived) pedagogical benefits and limitations of using ChatGPT in language education?***

- RQ1-A: *What are the main pedagogical benefits of using ChatGPT in language education, according to existing research?*
- RQ1-B: *What limitations related to ChatGPT integration have been identified in the current academic research?*

RQ3: ***How do (future) language teachers perceive and evaluate the quality of ChatGPT-generated content?***

- RQ3-C: *Do (future) language teachers use their critical thinking skills when using and evaluating the quality of ChatGPT?*

RQ4: ***What training do (future) language teachers receive or expect regarding ChatGPT?***

- RQ4-A: *Do (future) language teachers receive training on ChatGPT?*
- RQ4-B: *Do (future) language teachers want to receive training on ChatGPT?*
- RQ4-C: *And what do experts recommend training teachers to use ChatGPT effectively in language education?*

1.1 How artificial intelligence developed?

Artificial intelligence has developed through cycles of innovation, setbacks, and regulation, shaping its influence across multiple domains, including education, healthcare, and governance. John McCarthy and other academics first proposed the idea of AI as a system that could solve problems through symbolic reasoning at the Dartmouth Conference in 1956, marking the official start of the field (Toosi et al., 2021: 4). Large progress was made as a result of this early optimism, but two major periods of stagnation, referred to as "AI Winters," occurred in the 1970s and the late 1980s to early 1990s due to technological constraints and irrational expectations (Toosi et al., 2021: 5). However, AI made a strong comeback in the late 1990s and early 2000s, thanks to advances in machine learning, neural networks, and computational power. Important moments, like Deep Blue winning against Garry Kasparov in 1997 and AlphaGo beating human Go champions in 2016, demonstrated that AI was becoming better at solving structured problems and recognizing complex patterns (Toosi et al., 2021: 9–11). The creation of sophisticated language models like ChatGPT was made possible by these advancements. Since its introduction, ChatGPT has gone through multiple versions, each improving in text generation accuracy, reasoning capabilities, and contextual understanding. GPT-1 (2018) laid the foundation for AI-powered text generation; GPT-2 (2019) introduced enhanced fluency but raised concerns about disinformation; GPT-3 (2020) expanded to 175 billion parameters, improving coherence and linguistic versatility; and GPT-4 (2023) further refined general knowledge processing and advanced reasoning (Sarrion, 2023: 5).

Beyond education, AI is also changing other creative areas, like architecture, music, and cinema. In the field of architecture, AI helps make the design process faster and more creative. It uses tools that help analyse data and suggest better design layouts more quickly (Rafsanjani

& Nabizadeh, 2023: 3). Moreover, AI tools also support user-friendly and eco-friendly designs by adjusting things like lighting or choosing materials that fit specific needs (Rafsanjani & Nabizadeh, 2023: 3). Similarly, in cinema, AI has reshaped storytelling, production, and audience engagement. It helps writers by studying story structures and generating dialogue, and it makes editing and special effects more efficient (Pradeep et al., 2023: 2–5). AI also improves how audiences experience film by giving personalized suggestions based on their preferences (Pradeep et al., 2023: 5). In music and the visual arts, tools like Midjourney and DALL·E allow people to create content quickly, but they also create problems. These tools often use artists' work without permission, which can hurt artists financially and damage their reputations (Jiang et al., 2023: 368). Artists argue that while AI-generated images look good, they lack the personal meaning, culture, and emotion found in human art (Jiang et al., 2023: 368). These tools are also taking away jobs from artists, making it harder for them to earn a living. This shows the need for strong ethical rules and clear consent when using artists' work to train AI.

Indeed, despite some progress, AI still faces significant difficulties, particularly regarding bias, transparency, and governance. Large AI models, such as BERT and GPT-3, have shown gender and racial biases, raising concerns about the fairness and accountability of AI-based decision-making (Toosi et al., 2021: 13). In addition, deep learning models often work like "black boxes," meaning their decision-making is unclear and hard to understand, which creates problems for building trust and following regulations (Toosi et al., 2021: 13). In response, experts emphasize the need for explainable AI (XAI), which aims to make AI systems more transparent, understandable, and accountable (Toosi et al., 2021: 22).

However, beyond these socio-technical problems, AI governance and standardization remain key concerns. Organizations such as ISO (International Organization for Standardization) and IEEE (Institute of Electrical and Electronics Engineers) have attempted to establish global guidelines for AI. These are meant to ensure trust, interoperability[1], and accountability. Yet, progress is slowed by weak enforcement and a lack of global coordination between countries (Cihon, 2021: 10-12). Antoher issue is the so-called "race to the bottom" in AI regulation. In this race, countries compete for AI investment by lowering ethical and safety standards, which poses a serious threat to responsible AI use (Cihon, 2021: 7-8). To prevent another AI Winter

[1] Interoperability is defined as "the ability to work together with other systems or pieces of equipment" (Cambridge Dictionary, n.d.).

and ensure sustainable growth, experts stress the importance of interdisciplinary collaboration, increased participation in AI standardization efforts, and stronger regulatory frameworks (Cihon, 2021: 17-23). Countries such as the United States and China have already taken significant steps in AI governance, influencing the global landscape of AI policy and development (Cihon, 2021: 21-22). Nonetheless, these efforts must be aligned with ethical considerations and global best practices to avoid repeating past mistakes associated with earlier AI winters.

As Toosi et al. (2021) explain, exaggerated promises and unrealistic expectations have often led to overinvestment and later disappointment, which has slowed down AI research and development (2021: 4–7). To prevent such cycles, governance mechanisms should promote responsible research practices, aligning innovation with societal needs and ethical guidelines. International standards, such as rules and principles that define how AI systems should be designed, tested, and used responsibly, play a key role in this process. Cihon (2021) highlights that such guidelines can not only improve how AI systems work together but also help build a culture of safety and responsibility. This ensures that AI research and development remain ethical and beneficial for society (2021: 8).

Conclusion: AI has evolved through repeated phases of progress and setbacks, leaving a significant mark on various sectors such as education, architecture, the arts, and beyond. Despite its impressive potential, AI also raises critical concerns regarding ethic bias, transparency, and governance. Past experiences of overexpectation and disillusionment showed the need for a realistic and responsible approach. To ensure AI remains beneficial and sustainable, future progress must be guided by ethical principles and social values.

1.2 Artificial intelligence in education

AI has increasingly transformed education in the French-speaking Community of Belgium. Indeed, for the past two years, some effects have been observed in language learning, personalized instruction, and automated feedback systems. However, its integration raises important pedagogical, ethical, and governance concerns. Historically, AI in education has developed alongside broader technological advancements, reflecting a co-evolution between humans and digital tools (Thorne, 2024: 568). AI's presence in education dates back to the

1960s, when early rule-based systems[2] tried to provide structured learning experiences (Tahiru, 2021: 2). Over time, AI evolved trough the development of machine learning[3] and natural language processing[4] (NLP), leading to the rise of intelligent tutoring systems[5] (ITS) that personalize learning based on student responses (Tahiru, 2021: 3). AI-driven platforms can now automate grading, refine instructional strategies, and provide real-time feedback, making education more accessible and adaptive (Thorne, 2024: 570). However, as Generative AI (GenAI)[6] continues to self-improve and optimize its behaviour in learning environments, concerns regarding bias, academic integrity, and AI's autonomy in decision-making have emerged (Thorne, 2024: 570).

One of the primary risks associated with AI in education is its tendency to maintain its current functioning and defend its original objectives, which may impact how it interacts with students and educators (Omohundro, 2018: 5). AI systems, particularly GenAI like ChatGPT, operate based on predefined utility functions which are sets of objectives that determine how the AI prioritizes tasks (Omohundro, 2018: 4). In educational contexts, AI may prioritize engagement and efficiency over deeper learning, potentially encouraging surface-level understanding rather than critical thinking (Thorne, 2024: 570). This issue is compounded by the fact that AI is trained on human-generated data, which inherently carries biases (Thorne, 2024: 570). These biases can affect how AI responds, potentially reinforcing stereotypes, creating misleading representations in language learning, and causing unfairness in how students are assessed (Thorne, 2024: 570).

Without proper oversight and pedagogical adjustments, AI could unintentionally shape students' linguistic development based on flawed or biased data (Thorne, 2024: 570). Nevertheless, it is important to note that teachers and manuals carry biases too.

---

[2] A rule-based system in AI is defined as "a system that applies human-made rules to store, sort and manipulate data. In doing so, it mimics human intelligence. Rule based system in AI require a set of facts or source of data, and a set of rules for manipulating that data. These rules are sometimes referred to as 'If statements' as they tend to follow the line of 'IF X happens THEN do Y'". (Engati, n.d.)

[3] Machine learning is defined as "a subfield of artificial intelligence, which is broadly defined as the capability of a machine to imitate intelligent human behavior". (Mitsloan, n.d.)

[4] Natural language processing (NLP) is defined as "the ability of a computer program to understand human language as it's spoken and written -- referred to as natural language. It's a component of artificial intelligence" (Techtarget, n.d.)

[5] An intelligent tutoring system (ITS) is defined as "a computer system that aims to provide immediate and customized instruction or feedback to learners, usually without requiring intervention from a human teacher. ITSs have the common goal of enabling learning in a meaningful and effective manner by using a variety of computing technologies." (Encyclopedia, n.d.)

[6] Generative AI refers to "deep-learning models that can generate high-quality text, images, and other content based on the data they were trained on." (IBM, n.d.)

Additionally, AI's potential to facilitate cheating has raised concerns among educators. With GenAI tools capable of generating essays, answers, and translations instantly, students may rely on AI-generated content without fully engaging with the learning material (Thorne, 2024: 570). This over-reliance can hinder critical thinking, writing skills, and independent problem-solving, ultimately undermining the integrity of educational assessments (Thorne, 2024: 570).

Despite these challenges, AI presents significant benefits in education, especially in supporting personalised language learning, enhancing student engagement, and improving the quality of feedback. AI-driven tools can tailor instruction, provide real-time feedback, and adapt to individual student needs, which is particularly helpful for multilingual learners (Thorne, 2024: 570-571). Moreover, AI conversation tools allow students to practice communication in a second language, refine their grammar, receive feedback that fits the situation, improving their engagement and motivation (Thorne, 2024: 570). However, the author emphasizes that while AI can enhance education, it cannot replace human teachers, as effective learning depends on social interaction, cultural immersion, and interpersonal relationships, which AI lacks (Thorne, 2024: 571).

To make sure AI truly supports education, experts highlight the need for critical AI literacy, clear ethical guidelines, and balanced teaching approaches (Tahiru, 2021: 16). It is important that students learn how to think critically about AI-generated content, understand its limitations, and use it as a collaborative tool rather than relying on it blindly (Thorne, 2024: 570–571). As Thorne argues, language learning is fundamentally driven by the human relationships it enables, and the integration of GenAI must prioritize these social and communicative goals while addressing the ethical and practical challenges it poses (2024: 571). In the future, education systems need to find a balance between using AI's potential and protecting learning that promote real understanding and critical thinking, making sure that AI remains a complementary tool rather than an independent authority in education (Thorne, 2024: 572). This co-evolutionary approach, which couples human intelligence and creativity with GenAI tools, offers a promising pathway for creating language learning environments that are fair and that supports all learners equally (2024: 571-572).

Conclusion: AI is playing an important role in education in French-speaking Belgium, offering opportunities in areas such as language learning, personalized instruction, and real-time feedback. These developments reflect the co-evolution between human and digital

intelligence. While the integration of GenAI raises important concerns, it also presents possibilities for enhancing student engagement, supporting diverse learning needs, and improving educational practices. To fully benefit from these innovations, it is essential to promote critical AI literacy, equip educators with the tools to guide use, and maintain a focus on the human connections that drive meaningful learning. When used responsibly, AI can become a powerful ally in creating more inclusive, adaptive, and effective learning environments.

1.3 General overview of ChatGPT

According to the official website, ChatGPT is a sibling model to InstructGPT, which is trained to follow an instruction in a prompt and provide a detailed response (OpenAI, n.d). ChatGPT is a generative pre-trained transformer (GPT) model designed to understand and generate human-like text using natural language processing (NLP) techniques (Sarrion, 2023: 3). Built on a neural network architecture, which means, a structure of an AI system inspired by the human brain, ChatGPT processes input data through multiple layers of interconnected neurons, mimicking the way the human brain analyzes language patterns and word relationships (Sarrion, 2023: 10).

ChatGPT operates in two main phases: training and fine-tuning. During training, the model is exposed to billions of words from diverse sources, including books, news articles, and online conversations, allowing it to grasp the subtleties of natural language and produce appropriate and coherent responses (Sarrion, 2023: 13). This process relies on NLP techniques such as tokenization, which breaks text into smaller units; word embeddings, which convert words into numerical vectors for semantic analysis; and language modeling, which predicts word sequences to ensure fluency (Sarrion, 2023: 11). In addition, attention mechanisms play a critical role in refining responses, as ChatGPT focuses on the most relevant parts of a sentence to generate contextually appropriate text (Sarrion, 2023: 12). Once the model has completed its initial training, it can be further adapted to specific tasks by being exposed to more targeted datasets. This fine-tuning process makes it flexible and suitable for various applications, including education, customer service, content creation, and machine translation (Sarrion, 2023: 20).

1.3.1 GPT models

The following information is based on OpenAI'[7]s official website, though it has been reformulated. According to OpenAI, the application offers several advanced models, each designed for specific functions and differing in reasoning ability, processing speed, and input-output capabilities. Below is an overview of the available models:

1. GPT-4o

GPT-4o (where "o" stands for "omni") is OpenAI's most recent leading model. It is designed for a broad range of tasks that require complex reasoning and problem-solving. The model accepts both text and image inputs and produces text-based outputs, including structured responses. It is characterised by high accuracy, flexibility, and fast processing speed. However, it does not support audio input or output (OpenAI, n.d.).

2. O3-mini

The o3-mini model is a small model developed with a focus on cost-efficiency and speed. Despite its limited size, it retains effective reasoning capabilities and performs well in structured outputs, function calling, and batch processing tasks. The model is restricted to text-only inputs and outputs and does not support audio or image functionalities (OpenAI, n.d*).*

3. GPT-4.5 (Research Preview)

GPT-4.5 is OpenAI's largest and most capable model to date, though it remains in a research preview phase. It is designed for tasks involving deep reasoning, creative thinking and agentic planning. The model is particularly suited for open-ended applications such as writing, learning, and idea development. It demonstrates improved understanding of user intent and an expended knowledge base. At present, it does not support image or audio output and has limited availability (OpenAI, n.d).

4. GPT-4o-mini

GPT-4o-mini is a smaller, more affordable variant of GPT-4o, designed for focused tasks that require efficiency rather than high-level reasoning. It is suitable for fine-tuning and cost-effective deployment, as outptus from GPT-4o can be distilled into this model. While it

[7] The artificial intelligence company and research lab that created it.

maintains speed and practicality, it is limited to text-based outputs and demonstrates a moderate level of reasoning capability (OpenAI, n.d).

5. GPT-o1

The o1 model series is specifically designed for advanced reasoning and problem-solving. These models are trained using reinforcement learning techniques, allowing them to develop deep, logical thought processes before providing responses. However, they tend to be slower than other models due to their extensive internal reasoning steps and it does not suppport image nor audio output (OpenAI, n.d).

6. Sora

Sora is OpenAI's video generation model, designed to create high-resolution (1080p) videos up to 20 seconds long using text, image, and video inputs. The model can generate new content from text descriptions, animate still images, or extend existing videos while maintaining visual consistency.

Sora operates using a diffusion process, which starts with static noise and gradually refines it into a coherent video. This method ensures smooth motion and subject continuity, even if elements move in and out of the frame. Additionally, Sora incorporates OpenAI's recaptioning technique (originally used in DALL·E 3), allowing it to better understand and follow user instructions when generating videos. This model is still in development, that is why there are potential risks of misuse, and so, misleading content (OpenAI, n.d).

1.3.2 ChatGPT's advantages

ChatGPT's versatility makes it useful for various applications, including chatbots and virtual assistants. Chatbots are computer programs designed to simulate conversation with human users, often used to provide instant responses and assistance through messaging platforms. In these roles, ChatGPT powers automated customer service, marketing interactions, and e-commerce platforms (Sarrion, 2023: 17-18). It also plays a key role in machine translation, where its multilingual training enables more natural translations than traditional rule-based systems (Sarrion, 2023: 18-19). In content generation, businesses use ChatGPT to create product descriptions, blog posts, reports, and summaries, reducing costs while maintaining efficiency (Sarrion, 2023: 20). Moreover, ChatGPT is being used more and more to help people find information, as it allows users to ask questions in everyday language and get relevant, clearly organised answers in return (Sarrion, 2023: 22–24)

While these advantages highlight the broader potential of ChatGPT, the section 3.4.1 will focus specifically on the benefits it offers in the context of language education.

1.3.3 ChatGPT's limitations

According to the official website of ChatGPT, OpenAI explicitly acknowledges that the tool has several significant limitations. One of the most problematic is its tendency to generate responses that, while appearing plausible, are ultimately incorrect or nonsensical. This issue is particularly difficult to address for several reasons. During the reinforcement learning process, a feedback-based method where the model learns through trial and error by receiving positive or negative signals, there is no absolute source of truth on which the model can rely. Furthermore, increasing the model's level of caution during training would lead to a higher number of refusals to respond, including in cases where the model could have provided an accurate answer. Supervised training, where AI learns from human-provided examples with correct answers, further complicates the issue, as the most appropriate response is based on the model's internal knowledge rather than the knowledge of the human trainers (OpenAI, n.d). This challenge is also highlighted by Sarrion (2023), who emphasizes that one of the most critical limitations of ChatGPT is its lack of deep contextual understanding. The model generates responses by identifying statistical patterns rather than through genuine comprehension, which frequently results in inaccurate or misleading information (2023: 14-16).

Another documented limitation relates to ChatGPT's sensitivity to variations in input phrasing. Even minor changes to the wording of a question can produce significantly different responses. In some cases, the model might indicate that it does not know the answer, while a slightly rephrased version of the same question might receive an accurate response (OpenAI, n.d). ChatGPT's verbosity (the tendency to use more words than necessary) also constitutes a known issue. The model tends to produce responses that are excessively long and often repeats certain phrases. This tendency stems from biases introduced during training, as human trainers often prefer longer answers that appear more comprehensive (OpenAI, n.d). This issue aligns with findings from Thorne (2024), who demonstrated that ChatGPT frequently violates Grice's maxim of quantity, a principle from pragmatics that suggests that speakers should provide the right amount of information (not too much, not too little) to keep communication effective (2024: 570). ChatGPT frequently violates this maxim by offering more detail than necessary, which can reduce clarity and user satisfaction. In addition to these technical challenges,

ChatGPT typically does not seek clarification when faced with ambiguous questions. Rather than asking follow-up questions to confirm the user's intent, it tends to make assumptions, which can lead to inaccurate or irrelevant responses (OpenAI, n.d).

Finally, despite OpenAI's efforts to program ChatGPT to refuse harmful or inappropriate requests, the model occasionally produces responses containing harmful content or biased viewpoints (OpenAI, n.d). Sarrion (2023) also underlines this issue, noting that the model's responses sometimes reflect societal stereotypes or skewed perspectives (2023: 16). For example, the model has at times associated leadership roles with men and caregiving roles with women, reflecting gender biases present in the training data. This concern is supported by a UNESCO study[8] revealing that Large Language Models (LLMs), including GPT-3.5, exhibit significant gender biases. The study found that women were frequently associated with domestic roles and terms like "home", "family" and "children", while men were linked to terms such as "business", "executive", "salary", and "career" (UNESCO, n.d). To mitigate this, OpenAI has implemented a Moderation API designed to detect and block unsafe content, though the system remains imperfect and may generate both false positives and false negatives. OpenAI actively encourages user feedback to further refine these safety measures (OpenAI, n.d).

To make these challenges more accessible, best practices have been developed for optimizing ChatGPT's performance. Ensuring high-quality input data is crucial, which involves collecting diverse datasets, cleaning and preprocessing text, and verifying data accuracy to improve response reliability (Sarrion, 2023: 55-56). Strategies to minimize bias include using diverse training sources, applying bias detection algorithms, and raising awareness among developers about potential ethical risks (Sarrion, 2023: 57). ChatGPT's performance can also be evaluated using various techniques, such as perplexity measurements, human evaluation, and Turing tests[9], which assess the model's fluency, coherence, and ability to mimic human responses (Sarrion, 2023: 57-58).

[8] (UNESCO, n.d.)

[9] The Turing Test is described as "a deceptively simple method of determining whether a machine can demonstrate human intelligence: If a machine can engage in a conversation with a human without being detected as a machine, it has demonstrated human intelligence.
The Turing Test was proposed in a paper published in 1950 by mathematician and computing pioneer Alan Turing. It has become a fundamental motivator in the theory and development of Artificial Intelligence (AI)" (Investopedia, n.d.)

Conclusion: While ChatGPT demonstrates great advantages, it still faces some limitations that affect its reliability, accuracy and ethical use. These challenges include the generation of plausible but incorrect information, sensitivity to input phrasing, verbosity, lack of contextual understanding, and occasional reproduction of biased or harmful content. Despite OpenAI's ongoing efforts to address these issues through moderation tools, user feedback, and bias reduction strategies, the system remains imperfect. Understanding these limitations is essential for responsible use, particularly in educational and professional contexts. By applying practices such as refining input quality, promoting diverse training data, and using evaluation techniques, users can help enhance ChatGPT's performance while remaining conscious of its constraints.

1.3.4 How to write an effective prompt for ChatGPT

This section is based on tips shared by OpenAI about how to write better prompts for ChatGPT. The advice has been reformulated, and it can help users get clearer, more useful answers, especially when using ChatGPT for learning or teaching.

One approach is to give ChatGPT a specific role, such as "You are a language teacher". This helps shape the response for a certain audience or context. Another important tip is to place instructions at the beginning of the prompt and use delimiters like ### or """ to clearly separate different parts of the request. For example, instead of just asking for a summary, a better prompt would be "Summarize the following text in three bullet points: "" [text] "" (OpenAI, n.d).

Being specific and detailed also improves the quality of responses. Instead of saying, "Write a poem about OpenAI", a better prompt would be "Write a short, inspiring poem about OpenAI in the style of William Blake, focusing on the launch of DALL-E" (OpenAI, n.d). If the task is complex, it helps to break it down into steps (a method called chain-of-thought prompting) to understand and process the task more effectively. Another strategy is providing examples (known as few-shot prompting) to show the AI what kind of response is expected. If a simple prompt does not work as intended, refining it with examples or adjusting the instructions can improve accuracy.

It is also important to avoid vague words. Instead of saying, "Keep the response short", it is better to write "Give a concise answer in three to five sentences". Moreover, prompts should guide the AI toward the right behavior rather than just telling it what not to do. For instance,

instead of saying, “Do not ask for passwords”, a better prompt would be “If the user asks for a password, direct them to www.samplewebsite.com/help/faq instead” (OpenAI, n.d).

By following these best practices, users can get more accurate, structured, and relevant responses, making ChatGPT a more effective tool for learning and problem-solving.

Conclusion : In short, getting the best results from ChatGPT depends on how well the prompt is written. Using clear roles, detailed instructions, examples, and specific formations helps guide the tool to give accurate and useful answers. By avoiding vague language and learning how to structure prompts effectively, users can fully benefit from ChatGPT as a powerful tool.

1.4 ChatGPT in language education

RQ1: ***What are the (perceived) pedagogical benefits and limitations of using ChatGPT in language education?***
RQ4: ***What training do (future) language teachers receive or expect regarding ChatGPT?***

1.4.1 Educational advantages of using ChatGPT in language education

RQ1-A: *What are the main pedagogical benefits of using ChatGPT in language education, according to existing research?*

ChatGPT offers many benefits in language education, enhancing both teaching practices and learning experiences. For teachers, it is beneficial by assisting in lesson planning and generating creative ideas. Consequently, it can significantly reduce teachers’ workload and diversify their activities (van den Berg & du Plessis, 2023: 10). Ghafouri et al. (2024) showed that ChatGPT could reduce stress and burnout by increasing teachers’ self-efficacy (2024: 17). Indeed, it could reduce their workload, save their time and energy, and make them more organised by asking the AI tool to design lesson plans, to determine course goals or to provide automatic feedback to students’ assessments (2024: 16-17).

For students, ChatGPT provides immediate feedback, particularly useful for writing assessments. Guo and Wang (2023) found out that ChatGPT’s feedback focuses on content, organisation and language, and that it is more directive than teachers’ feedback. This is why it can be argued that complementing teachers’ feedback with AI’s can improve students’ writing

by receiving more diverse comments. This enriched feedback process benefits both students and teachers by providing a more comprehensive perspective on writing improvement (2023: 8450-8452). For example, Zeevy Solovey (2024) showed that ChatGPT provides explanations, specific corrections or alternatives to students' work, which, with teacher guidance, can help students build autonomy by critically engaging with AI-generated feedback (2024: 12-16).

ChatGPT is a valuable tool to enhance students' autonomy and engagement too. According to Du and Alm (2024), it provides a user-friendly, non-judgemental environment which ables the students to express themselves and reduce their language anxiety (2024: 13-14). Indeed, ChatGPT creates opportunities for more interactive and engaging learning experiences by simulating real-life conversations. Teachers noted that students could practice communication skills by posing questions to ChatGPT and receiving natural language responses, enabling them to develop conversational fluency in English (Annamalai, 2024: 14). In addition, as Annamalai (2024) explains, teachers participating in the study identified that traditional classroom instruction often follows a "one-size-fits-all" approach, which fails to accommodate the diverse needs and learning preferences of individual students (2024: 7). In contrast, ChatGPT offers adaptive responses and tailored feedback, allowing students to receive individualised support based on their specific language proficiency levels (2024: 8). For example, students can input sentences into ChatGPT and receive instant corrections related to grammar, vocabulary, and spelling (2024: 8). This type of immediate, personalised feedback is particularly valuable for fostering independent learning, as it allows students to monitor their progress and identify areas for improvement without having to wait for teacher feedback. Moreover, it can assist students in generating ideas, making it easier for them start writing and express their creativity (2024: 14). Ahn et al. (2024) agree with Annamalai (2024) on ChatGPT's potential to improve student's autonomy because the AI tool can adapt its output to the learners' needs and create activities based on their proficiency (2024: 351). Juan et al. (2023) showed that students perceived an improvement in their motivation by using ChatGPT because it can provide various learning materials and exercises, and adapt the content and difficulty to the students' level (2023: 4065). Furthermore, ChatGPT's 24/7 availability allows students to access personalised learning outside the classroom, promoting self-directed learning and autonomy (Annamalai, 2024: 9). This flexibility is particularly advantageous for students who need additional practice or those who prefer to learn at their own pace. Thus, by enabling personalised, interactive, and flexible learning, ChatGPT has the potential to increase student motivation and engagement in English language education.

In conclusion, ChatGPT presents significant advantages for both teachers and students in language education. For teachers, the tool reduces workload, supports creative lesson planning, and enhances efficiency, ultimately contributing to lower stress and burnout. For learners, it offers immediate, personalised and varied feedback, particularly in writing skills. Its ability to adapt responses to each student's level helps support autonomy, lower language anxiety, and increase motivation and engagement. In addition, its easy access allows students to learn independently and at their own pace, outside the classroom. As a result, using ChatGPT as a complementary tool can improve teaching and make the learning experience more effective.

1.4.2 Risks, limitations and challenges of using ChatGPT in language education

RQ1-B: *What limitations related to ChatGPT integration have been identified in the current academic research?*

While ChatGPT offers valuable benefits in language education, it also presents a wide range of ethical, social and pedagogical challenges that educators and institutions must address thoughtfully.

From an ethical perspective, one of the primary concerns is the potential for academic dishonesty and plagiarism. Ulla et al. (2023) highlighted that some teachers fear students might use ChatGPT for copy-pasting answers instead of engaging with the material in a thoughtful way. This issue is linked to another concern: the lack of critical thinking. Students may rely on this tool without exercising their judgment, which could prevent them from improving their language skills (2023: 176). This concern extends beyond students as Juan et al. (2023) warned that teachers themselves could over-rely on ChatGPT, incorporating AI-generated content without sufficient evaluation, which increases the risk of passing inaccurate or irrelevant information to students (2023: 4065). Moreover, the issue of critical thinking is closely tied to these ethical concerns. When students heavily depend on ChatGPT for answers, they may fail to develop analytical skills, becoming passive recipients rather than active learners (Al-Khresheh, 2024: 9).

In addition to ethical issues, ChatGPT raises serious social concerns related to classroom dynamics and relationships. Ghafouri et al. (2024) identified that the introduction of ChatGPT could alter the authority balance between teachers and students, as well as among students

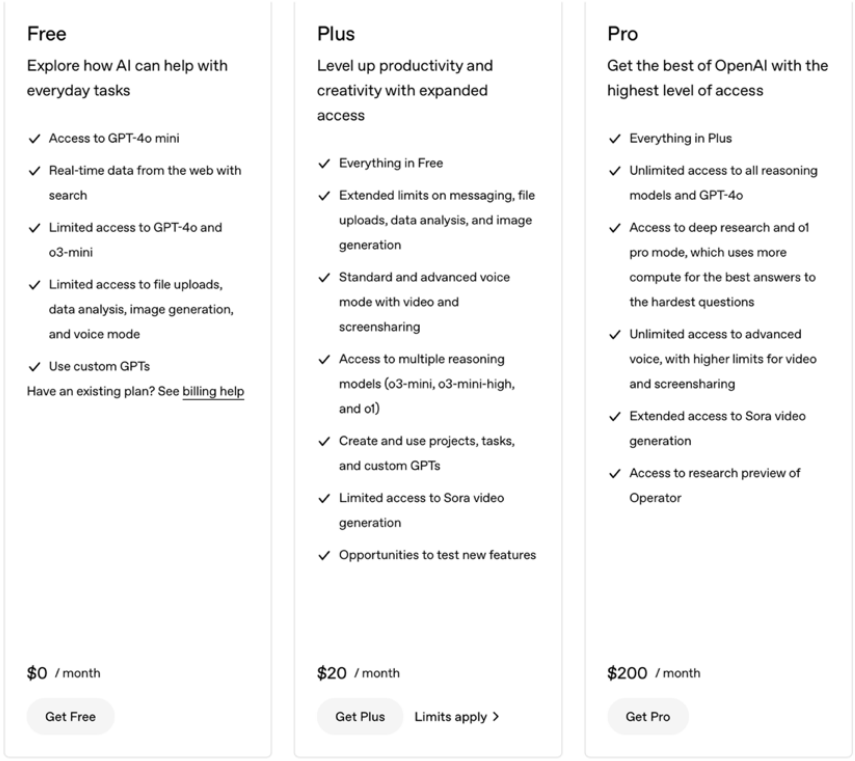
themselves. When learners turn to ChatGPT for answers instead of engaging with their teachers or peers, this shift can weaken the teacher-student relationship and reduce meaningful peer interaction (2024: 18). Furthermore, some educators fear that ChatGPT's convenience might lead to decreased classroom attendance, as students may perceive they can learn just as effectively by relying on AI at home (Ulla et al., 2023: 176). This shift in classroom habits risks diminishing collaborative learning opportunities, which are essential for language acquisition. In multicultural classrooms, ChatGPT's lack of cultural and emotional sensitivity exacerbates these social challenges. Nizzolino (2024) stresses that ChatGPT lacks the cultural awareness and empathy that human teachers should naturally bring to the classroom, which means it may unintentionally generate responses that feel inappropriate or culturally insensitive (2024: 258). This absence of cultural and emotional understanding can lead to frustration, disengagement, and ultimately a decline in student motivation, especially in linguistically and culturally diverse settings.

Moreover, unequal access to technology raises further equity concerns. Yuan et al. (2024) pointed out that not all students have equal access to ChatGPT, particularly when premium versions offer more advanced capabilities. This discrepancy could exacerbate existing inequalities in educational outcomes (2024: 51).

According to OpenAI's official website, the free version gives full access to the GPT-4o mini model, but access to GPT-4o and GPT-3-mini is more limited. Features alike file uploads, data analysis, image generation, and voice messaging are also restricted. However, this version can access real-time information from the internet and remembers earlier inputs during a conversation (OpenAI, n.d).

The Plus version, which costs $20 per month, gives access to more models, including GPT-3-mini, GPT-3-mini-high, GPT-1, and limited access to Sora for video generation. This version supports greater productivity and creativity, although some restrictions still apply, but fewer than in the free version. It also includes both standard and advanced voice features, along with video and screen sharing (OpenAI, n.d).

The Pro version, priced at $200 per month, offers full access to all models, including GPT-4o, and extended access to Sora video generation. It is suited for advanced research and includes early access to the Operator research preview. Pro users also get unlimited use of both standard and advanced voice models (OpenAI, n.d).



Screenshot 1: Overview of the different ChatGPT subscriptions available on the OpenAI platform.

This table highlights the unequal distribution of access to ChatGPT's most advanced features which are reserved for paying users. As a result, students and teachers who cannot afford premium subscriptions may be limited to less capable versions, reinforcing existing educational disparities and restricting equal learning and teaching opportunities.

In addition, ChatGPT presents significant challenges related to its effective use in instructional design and classroom implementation. Ghafouri et al. (2024) emphasised that ensuring ChatGPT's outputs align with curriculum goals, learning objectives, and assessment criteria is essential (2024: 18), but this alignment requires significant teacher effort. In practice, ChatGPT's responses often need to be rephrased, adapted, or contextualised to suit specific classroom needs, which can become time-consuming (Solovey, 2024: 14). Moreover, effective use of ChatGPT heavily depends on teachers' ability to formulate precise prompts and critically evaluate AI-generated content (Baskara, 2023: 265). However, teachers should receive sufficient training in these areas, in order to avoid either underuse or misuse of the tool (Al-Khresheh, 2024: 9). Ahn et al. (2024) further highlight that poorly formulated prompts frequently lead to vague, inaccurate, or irrelevant responses, limiting ChatGPT's educational value (2024: 352). These technical and pedagogical challenges contribute to a broader concern about over-reliance. When students and teachers rely too heavily on ChatGPT, both risk losing

opportunities to develop independent thinking, problem-solving, and creative expression (Ulla et al., 2023: 178). This over-reliance can also make students less able to handle difficult language tasks that need more advanced thinking.

Finally, ChatGPT's potential for manipulation and undue influence adds another layer of complexity. Nizzolino (2024) warns that ChatGPT's persuasive capabilities, combined with its lack of transparency about the sources and reasoning behind its responses, could make students vulnerable to subtle manipulation or the spread of biased information (2024: 258). This is particularly worrying in contexts where students may not have the critical media literacy needed to question or verify AI-generated content. Furthermore, because ChatGPT's responses reflect the biases present in its training data, students could be exposed to content that reinforces harmful stereotypes related to gender, culture, or socioeconomic status (see page 109) if prompts are unclear or ambiguous (Kotska & Toncelli, 2023: 7-13).

Conclusion: While ChatGPT offers support in language education, its use also presents ethical, social, and pedagogical challenges that must be carefully managed. Concerns such as academic dishonesty, weakened critical thinking, altered classroom dynamics, and reduced cultural sensitivity highlight the need for thoughtful integration. Moreover, unequal access to advanced features risks reinforcing educational inequalities. Effective implementation requires teacher training, structured teaching approach, and critical evaluation of AI outputs. Without these protections, there is a risk of over-reliance, reduced learning quality, and exposure to biased or misleading content.

1.4.3 Pedagogical suggestions for integrating ChatGPT in language education

RQ4-C: *And what do experts recommend training teachers to use ChatGPT effectively in language education?*

The following pedagogical suggestions are grounded in recent academic research, which emphasises the need for thoughtful integration of ChatGPT to support, rather than replace, traditional language teaching practices.

First and foremost, teachers should be equipped with targeted training on prompt engineering and AI literacy. By understanding how to design specific and relevant prompts, teachers can

guide ChatGPT's responses to align with educational objectives and generate accurate content that supports their lesson plans and students needs. However, educators should remain critical of its responses and verify the generated content (Urazbayeva et al., 2024: 13).

Moreover, it is essential to establish clear guidelines on responsible AI use in the classroom. Teachers should inform students about ChatGPT's limitations, emphasising the need for critical evaluation of AI-generated content. Teachers could create activities to foster an environnement where students are encouraged to think critically rather than accept AI outputs automatically (Chan & Lee, 2023: 18).

In addition, integrating ChatGPT into collaborative learning activities can further enhance its educational value. For instance, students can work in pairs or small groups to evaluate and discuss AI-generated content, identifying strengths and weaknesses in ChatGPT's responses. This approach promotes peer-to-peer interaction and enhances communication skills, creating a balanced blend of AI support and human engagement (Du & Alm, 2024: 13). However, this kind of activity may be difficult to organise in secondary schools.

Lastly, teachers should adopt a reflective approach to monitor ChatGPT's impact on learning. Al-Khresheh (2024) stresses the need for teachers to adapt their curriculum, watching how well ChatGPT helps with learning goals and adjusting their methods as needed (2024: 10). Similarly, Ahn et al. (2024) insist on the importance of understanding that ChatGPT cannot replace a teacher, but it can assist him (2024: 354).

Conclusion: The effective integration of ChatGPT into language education requires thoughtful teacher training, clear classroom guidelines, and a reflective approach to designing lessons. Teachers must critically evaluate AI-generated content, guide students in responsible use, and promote collaborative activities that balance AI support with human interaction.

1.4.4 Managing critical thinking for students when using ChatGPT

For university students, managing critical thinking when using ChatGPT requires actively engaging with AI-generated content, rather than passively accepting the responses provided by the model (Suriano et al., 2024: 2). According to Suriano et al. (2024), students must question, analyze, and cross-check ChatGPT's responses with other reliable sources, ensuring they develop strong reasoning and evaluation skills instead of treating AI content as inherently

credible (2024: 2). This approach is closely related to Cognitive Load Theory, which explains how the brain processes information. The theory suggests that learners have a limited capacity for handling information at one time, and if too much is provided without active processing, it can lead to shallow understanding. While ChatGPT can reduce cognitive load by summarizing or organizing content, over-reliance on the tool may limit deeper mental engagement, causing students to focus only on surface-level information rather than developing higher-order thinking skills (Suriano et al., 2024: 2). To address this, educators should explicitly teach students how to critically interrogate ChatGPT outputs, guiding them to evaluate reliability, detect bias, and assess the appropriateness of AI-generated responses for academic purposes (Suriano et al., 2024: 7). Additionally, integrating critical media literacy into university courses can equip students with the necessary tools to understand the socio-political impacts of AI (Trust et al., 2023: 11). By combining lessons on critical thinking, active engagement, and clear guidelines on responsible AI use, students can develop the skills to use ChatGPT productively without losing their ability to think critically (Suriano et al., 2024: 7).

For high school students, to ensure that ChatGPT contributes positively to students' development of critical thinking, educators must also provide structured learning activities that actively engage students in questioning, analyzing, and validating the information generated by the tool. Guo and Lee (2023) introduced a three-stage ChatGPT activity in introductory chemistry courses, where students first familiarized themselves with the tool, then created essays using ChatGPT, and finally reviewed and validated the AI's output (2023: 4877). This process significantly increased students' confidence in asking insightful questions, analyzing content critically, and drawing logical conclusions (2023: 4880). However, the study also showed that some students struggled to check if ChatGPT's references were reliable, especially when the AI gave fake sources or sources that are difficult to verify. This pushed them to improve their ability to double-check information using trusted academic sources (2023: 4881).

Conclusion: Helping students use ChatGPT in a critical way means they must actively engage with its content. University students need support to question, verify, and analyse AI responses, along with courses in critical media literacy. For high school students, guided activities like reviewing AI-generated work can build confidence and thinking skills, though challenges such as veryfying sources can still be difficult. With proper support, ChatGPT can enhance learning while maintaining students' critical thinking.

1.4.5 Managing critical thinking for teachers when using ChatGPT

RQ3-C: *Do (future) language teachers use their critical thinking skills when using and evaluating the quality of ChatGPT?*

For teachers, managing critical thinking when using ChatGPT centers on critical evaluation, adaptation, and collaborative reflection, especially when using AI-generated content for lesson planning or classroom instruction (van den Berg & du Plessis, 2023: 1). ChatGPT can provide efficient lesson plan templates, activity suggestions, and teaching materials, but teachers must actively evaluate, adapt, and contextualize this content to suit the needs of their specific learners and teaching environments (van den Berg & du Plessis, 2023: 7). This critical adaptation process encourages thinking, prompting teachers to reflect on accuracy, creativity, and cultural or contextual relevance, rather than simply adopting AI suggestions without modification (van den Berg & du Plessis, 2023: 10). However, relying too much on ChatGPT could weaken teachers' creativity and critical thinking, especially if they start letting AI make decisions for them, such as letting it determine how lessons should be structured or delivered (van den Berg & du Plessis, 2023: 8). To prevent this, teacher education programs should actively train future teachers not only in how to use ChatGPT effectively, but also in how to assess and refine AI-generated content critically, fostering a reflective teaching practice that integrates AI as a supporting tool rather than a decision-maker (van den Berg & du Plessis 2023: 8). In addition, Trust et al. (2023) emphasize the importance of embedding critical AI literacy into teacher education, so that teachers themselves become role models of responsible AI use in the classroom (2023: 14). This includes not only evaluating content quality but also engaging in transparent conversations with students about the limitations, biases, and ethical issues related to AI use in education (Trust et al., 2023: 11). Finally, Open Educational Resources (OER) produced through AI tools like ChatGPT may improve access to teaching materials, but their quality is not guaranteed, making critical evaluation skills indispensable for teachers who rely on these resources (van den Berg & du Plessis 2023: 10). By encouraging critical reflection, providing clear training on evaluation, and transparent dialogue about AI's role, teachers can effectively manage ChatGPT's integration into teaching while protecting their own critical thinking skills.

Conclusion: While ChatGPT can support lesson planning and teaching, its effective use in education depends on teachers' ability to critically evaluate and adapt AI-generated content. Without thoughtful engagement, there is a risk of diminishing teahers' creativity and professional judgement. To prevent this, teacher training must prioritise the development of critical AI literacy, reflective thinking, and ethical awareness. With these skills, educators could evaluate content quality, relevance, and bias, allowing ChatGPT to be used as a helpful tool.

1.4.6 ChatGPT and equity

Using ChatGPT in education brings both opportunities and challenges in terms of equity and inclusion. That is why it is essential for educational institutions to adopt critical and proactive strategies to ensure fair access and benefits for all learners. According to García-López et al. (2024), one of the most promising aspects of ChatGPT is its ability to personalize learning, allowing students to receive adaptive support adjusted to their individual needs and preferences, which could particularly benefit students with disabilities or those facing economic or geographic barriers, such as living in remote areas with limited internet access or educational resources (2024: 2-8). This personalization, however, comes with privacy risks, as large amounts of personal data must be collected, processed, and analyzed to enable this adaptive learning (2024: 2). Similarly, Rawas (2024) points out that ChatGPT can help make higher education more equal by offering personalized learning paths that match each student's learning style, background knowledge, and individual needs (2024: 6899). However, both authors emphasize that this potential will only be realized if equity challenges, including technological access and algorithmic bias[10], are actively addressed (García-López et al., 2024: 10-11; Rawas, 2024: 6901-6902).

Algorithmic bias is a major concern because ChatGPT is trained on vast datasets that often reflect existing cultural, linguistic, and social biases, which can lead to discriminatory content generation (García-López et al., 2024: 8; Rawas, 2024: 6902). As Rawas (2024) warns, this bias could make students from minority or marginalized groups feel left out because the AI might not recognize their experiences or cultures. This could lower their engagement and make

[10] "Algorithmic bias occurs when systematic errors in machine learning algorithms produce unfair or discriminatory outcomes. It often reflects or reinforces existing socioeconomic, racial and gender biases." (IBM, n.d.)

them feel even more excluded (2024: 6902). An example of that could be ChatGPT unintentionally prioritizing Christian holidays, for example, over other celebrations.

Furthermore, García-López et al. (2024) argue that ensuring equitable access to necessary technology, as well as providing training and support for all students, is essential to prevent AI from becoming a tool that only benefits the privileged (2024: 10).

Both articles emphasise that schools and universities need clear policies and ethical guidelines to make sure AI is used fairly (García-López et al., 2024: 2; Rawas, 2024: 6903). This means checking AI systems for fairness, using transparent algorithms, and keeping humans involved to catch and fix any biased or harmful results (García-López et al., 2024: 2). Institutions should also be open about how they collect, store, and use data, so that students know what is happening and their privacy rights are protected (García-López et al., 2024: 8; Rawas, 2024: 6902). Finally, training both teachers and students to use ChatGPT critically and effectively is essential to ensuring that everybody, regardless of socioeconomic background, cultural identity, or prior digital experience, can benefit from AI learning (García-López et al., 2024: 8; Rawas, 2024: 6905).

Conclusion: Making ChatGPT fair for everyone means having a clear plan that includes fair technology, strong privacy rules, equal access to digital tools, and helping people build digital skills, so that AI helps all learners instead of making existing gaps worse.

1.4.7 ChatGPT and assessment: what teachers can do

As ChatGPT becomes more present in language education, it raises new questions about how we assess students. This shift encourages educators to revisit traditional evaluation methods to protect academic integrity while taking advantage of AI's potential. According to Koraishi (2023), ChatGPT can enhance formative and summative assessments by providing instant, personalized feedback, and alignment with standardized language proficiency frameworks, such as the CEFR (2023: 65-67). The tool can also assist in diagnosing students' knowledge and tracking their progress over time (Koraishi, 2023: 65).

However, the use of GenAI in education also raises concerns. A major challenge highlighted by Thorne is cheating. He claims that "Students may (and do) misuse GenAI to generate plagiarized content without understanding—or even substantively addressing—the subject

matter, undermining the integrity of assessments and devaluing the educational process" (Thorne, 2024: 570). This issue is compounded by the limitations of traditional plagiarism detection tools. As demonstrated in an experiment by Stutz et al. (2023), traditional plagiarism checkers, such as Turnitin, failed to detect AI-generated work, whereas AI detectors, such as ZeroGPT, successfully identified it (2023: 5). This suggests that educational institutions must adopt AI-specific detection tools to maintain assessment credibility (Stutz et al. 2023: 6). Nonetheless, it is important to clarify that ZeroGPT is also an AI-based tool, and it can produce erroneous results too, indicating the need for cautious use.

In addition, a range of AI detection tools has been developed in response to the increasing use of AI-generated content. According to Elkhatat et al. (2023), some of the most commonly used detection tools include OpenOpen AI Classifier, Writer, Copyleaks, GPTZero, and CrossPlag, each of which analyzes linguistic patterns, sentence structure, and probability distributions to determine the likelihood that a text was generated by AI (2023: 3). Similarly, Malik and Amjad (2025) evaluated the accuracy of Turnitin, ZeroGPT, GPTZero, and Writer AI, particularly when detecting outputs from ChatGPT, Perplexity, and Gemini (2025: 91). Their study found Turnitin to be the most accurate, even when texts had been paraphrased or human-edited, while Writer AI was the least effective (Malik & Amjad, 2025: 96).

Despite these efforts, false positives remain a concern. Elkhatat et al. (2023) warn that students' original work may be incorrectly flagged as AI-generated, potentially leading to unjust penalties (2023: 8). The authors also observed that detection tools were more successful in identifying GPT-3.5-generated content than GPT-4, which is more fluent and difficult to classify (2023: 8). Dalalah and Dalalah (2023) further highlight that detection accuracy varies across different sections of academic writing, with literature reviews being more frequently flagged as AI-generated, likely due to their formal and widely cited nature. This inconsistency undermines trust in detection software, as it may misidentify legitimate research writing as AI-generated content (2023: 12). Moreover, paraphrasing tools like Quillbot and grammar checkers like Grammarly significantly reduce the effectiveness of AI detection tools, further complicating their reliability (Malik & Amjad, 2025: 97).

To address these limitations, several researchers suggest combining detection software with human review. Elkhatat et al. (2023) advocate a hybrid approach to improve reliability and reduce false positives (2023: 14). Dalalah and Dalalah (2023) emphasise the need for

institutions to establish clear policies regarding AI use and to ensure that detection tools are used thoughtfully rather than as unquestioned measures of academic integrity (2023: 11).

Alongside detection, scholars also propose adapting evaluation methods. Ghafar (2023) recommends restructuring evaluation by incorporating in-class writing tasks, oral presentations, and multimedia projects, which require students to demonstrate their language proficiency. Moreover, educators should not forbid using technology but rather integrate it into assignments, as this reflects how language is used in real-life situations. He also suggests integrating technology into assignments rather than prohibiting its use, as this better reflects real-world language use (2023: 79). Stutz et al. (2023) similarly argue that incorporating oral discussions and explanations as part of the evaluation process makes it harder for students to submit AI-generated work undetected (2023: 7). By shifting evaluation toward student reasoning and problem-solving, educators can help maintain academic integrity while adapting to the use of AI in education (Stutz et al. 2023: 7). Ghafar (2023) supports this idea and suggests using flipped classroom and ongoing assessments, which focus more on the learning process than on final exams helped by AI (2023: 79-80). However, it is important to note that ChatGPT is also trained for problem-solving tasks, which may pose additional challenges when designing such assessments.

Conclusion: While ChatGPT offers new possibilities for language assessment, it also raises critical concerns about academic integrity and detection reliability. To make the most of its potential while protecting academic integrity, teachers and institutions are encouraged to rethink their assessment practices, use reliable AI detection tools, and develop clear thoughtful policies around AI use in the classroom.

1.5 Artificial intelligence in teacher training at the University of Liège

RQ4-A: *Do (future) language teachers receive training on ChatGPT?*
RQ4-B: *Do (future) language teachers want to receive training on ChatGPT?*

The integration of AI in language education has become an important topic in the initial teacher training programs at the University of Liège. Over the past year, several initiatives, conferences, and workshops have been organized to explore the role of AI in teaching and learning in the field of language education.

1.5.1 Artificial intelligence in the “Didactics of Modern Languages” Course

As part of the “Didactics of Modern Languages” course, students had the opportunity to attend presentations on the use of digital tools and AI in language teaching. Two experts were invited to share their insights:

- Tiziana Paganelli gave a presentation titled “Choix et intégration des outils numériques à ma pratique pédagogique en langues modernes”. In her talk, she discussed various digital tools available for language teaching and briefly addressed the use of ChatGPT in the classroom.
- Audrey Thonnard presented a module focusing specifically on ChatGPT, providing a more in-depth exploration of its potential applications, benefits, and challenges in language teaching.

These interventions introduced future teachers to practical ways of integrating AI tools into their didactic sequences.

1.5.2 Artificial intelligence in conferences and research at the Institute of Languages in Liège

Beyond the “Didactics of Languages” course, the ISLV (*Insitut Supérieur des Langues Vivantes*) at the University of Liège has actively engaged in discussions on AI in education. Several conferences, research groups, and documents have been dedicated to exploring the pedagogical implications of AI tools such as ChatGPT:

- “A Teacher’s Prompt Guide to ChatGPT aligned with ‘What Works Best’”: This document, led by @herfteducator, provided teachers with strategies for effectively using ChatGPT in their practice, ensuring that AI integration aligns with evidence-based teaching principles.
- Working Group on “Intelligences artificielles génératives et enseignement/apprentissage des langues étrangères”: This initiative brought together researchers and educators to examine how generative AI can support language learning and teaching methodologies.
- “La perspective des étudiants sur l’utilisation de l’IA dans l’enseignement en FSA”: This document reflects students’ perspectives on AI integration in education, offering valuable insights into how learners perceive and interact with these technologies.

1.5.3 Steven Thorne and the Chaire Francqui

A major event in AI and language education at the University of Liège took place in February 2025, when Professor Steven Thorne visited as part of the Chaire Francqui program. Steven Thorne is a leading researcher in applied linguistics and language education, particularly known for his work on digital technologies, language learning, and intercultural communication. His research explores how digital tools shape second language acquisition. During his visit, he delivered conferences in Liège sharing his expertise on ChatGPT in language education. His presence was part of a broader academic initiative, as he is spending six months at the Université Catholique de Louvain (UCLouvain) under the Chaire Francqui program. His insights provided valuable perspectives on how AI can enhance language learning, while also highlighting potential risks and ethical considerations.

1.5.4 Conclusion

The University of Liège actively promotes the integration of AI into teacher training, helping future educators understand and use these tools in their teaching. Through guest lectures, research initiatives, official publications, and international partnerships, both students and staff are developing a solid understanding of AI in language education, especially regarding ChatGPT and generative AI.

This growing focus on AI highlights the importance of building AI literacy and learning how to use tools like ChatGPT effectively. In this context, my dissertation contributes to the ongoing conversation by exploring ChatGPT's potential in English language teaching, assessing its strengths and limitations, and showing how it can be used to enhance didactic sequences and help novice teachers.

1.6 Artificial intelligence in ongoing teacher training in Wallonia-Brussels Federation

In the Wallonia-Brussels Federation, professional development in digital education and AI has become increasingly central to teacher training policies. Several public platforms such as the IFPC (*Institut de la Formation Professionnelle en cours de Carrière*), EduLAB, and Digital Wallonia now provide a growing number of training sessions to help educators adapt to the evolving technology.

The IFPC, as the main provider of in-service training, offers over 500 courses annually. Among these, several sessions specifically target digital tools and AI. Notable examples include “Intelligence artificielle : renforcer l’expérience d’apprentissage grâce à l’IA générative (notamment ChatGPT)”, a practical session designed to guide teachers in integrating generative AI into their lessons. This training was delivered multiple times in early 2025, both for teachers and for school leadership. Another relevant session, “Formation à distance : Prise en main avancée et techniques professionnelles pour la recherche d’information, la veille et l’usage des ressources du Web (y compris l’intelligence artificielle) à intégrer dans vos projets pédagogiques”, focuses on developing advanced web-based pedagogical practices, including the use of AI tools.

Additionally, the IFPC promotes hands-on sessions exploring user-friendly educational tools. Examples include “Booster l’interactivité avec Genially” (April 2025), “LearningApps : un outil en ligne pour dynamiser ses enseignements” (April 2025), and “La Digitale – petits outils numériques simples et gratuits” (February 2025), all aimed at helping teachers create more interactive and student-centered language lessons.

Beyond individual tools, broader pedagogical integration is covered in the course “Intégrer les outils numériques dans sa pratique quotidienne d’enseignant” (April 2025), which helps teachers build concrete digital lesson plans across disciplines. Sessions like “Évaluer avec les outils numériques” also reflect the growing interest in digital assessment strategies, although some events (e.g., March 2025) have faced cancellations due to limited enrollment or logistical constraints.

EduLAB complements these offerings with the “Parcours Référent IA”, a modular training path launched in 2024. Teachers enrolled in this program select several AI-focused workshops, such as “Créer des contenus pédagogiques avec ChatGPT”, and are awarded a digital badge upon completion. This initiative is supported by the WBF and recognized by the IFPC, with new sessions planned into 2026.

The AI 4 Edu platform, developed by Digital Wallonia, acts as a central resource for teacher training in artificial intelligence, offering webinars, toolkits, and on-demand content. A recent example is the workshop “ChatGPT : Révolutionnez vos cours et vos techniques

pédagogiques", delivered in May 2024, which introduced teachers to prompt engineering and lesson planning with generative AI.

Lastly, major educational conferences such as Ludovia#BE and SETT provide additional training opportunities through conferences and practical workshops, often centered on language teaching, digital tools, and classroom applications of AI. These events are generally listed in the IFPC's approved training catalogue and allow teachers to stay informed about the latest technological and pedagogical developments.

Altogether, these initiatives demonstrate a clear commitment to equipping language teachers in the WBF with the knowledge and tools necessary to navigate digital transformation and responsibly integrate AI into their educational practices.

Conclusion: The WBF actively supports language teachers who want to explore digital tools and artificial intelligence in their practice. Through initiatives like the IFPC, EduLAB, and Digital Wallonia, teachers have access to practical, targeted, and up-to-date training opportunities. These programmes do not just help educators use tools like ChatGPT in their classrooms, they also encourage thoughtful reflection and alignment with pedagogical principles, ensuring technology is used in a responsible way.

2. Identification of the main characteristics of the current foreign language prescriptions in the Wallonia-Brussels Federation and of the didactic tools likely to fill the gaps in those prescriptions

2.1 Introduction

To assess whether ChatGPT can generate a didactic sequence that aligns with the educational standards of the French-speaking Community of Belgium, I first needed to isolate the main elements found in the official documents. These include the communicative and action-oriented approach, the balanced integration of the four language skills (listening, reading, writing, oral production with and without interaction), the CEFR levels, the Unités d'Acquis d'Apprentissage (UAAs), the "Connaître–Appliquer–Transférer" triplet, the inductive approach of grammar, and the use of learning strategies.

As a future teacher, I found that although these documents clearly define what is expected, they offer little support on how to apply these requirements in practice. That is why I turned to the "Didactics of Modern Languages" course taught by Mr Simons, which provided helpful tools for translating theory into classroom practice. This course introduced the Presentation–Fixation–Exploitation (PFE) model, which I chose to use as the main structure for analysing didactic sequences. While not mentioned in the official documents, it offers a practical structure for planning and evaluating lessons. In parallel, the action-oriented approach was explored through the use of textual genres, which I also included as a criterion in my evaluation.

All these elements taken from the official documents and the complementary training were integrated into the evaluation grids I developed to assess the sequences generated by ChatGPT. These grids are mainly based on the PFE approach but also include key aspects such as metacognitive strategies, the inductive approach, and the integration of textual genres. Finally, this chapter also aims to examine whether digital tools or artificial intelligence, including ChatGPT, are mentioned in the current official prescriptions.

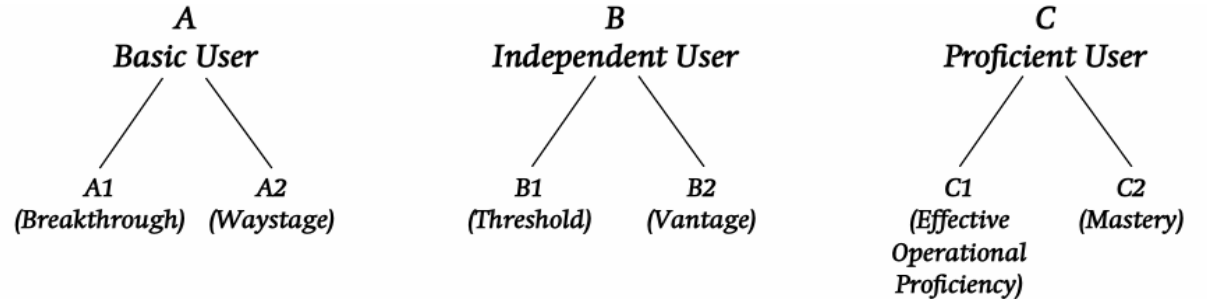
2.2 The Common European Framework of Reference for Languages[11]

2.2.1 Proficiency levels and impact in Belgian education

At the heart of the CEFR lies a six-level proficiency scale, ranging from A1 (Breakthrough) to C2 (Mastery), grouped into three broad categories: Basic, Independent, and Proficient User. These levels are defined in terms of what learners can do with the language in authentic communicative contexts. They apply across the four skills (listening, reading, speaking, and writing) and are used to assess both global and specific language abilities (Council of Europe, 2001: 23–30). These descriptors are widely adopted in Europe and guide both formative and summative assessments (2001: 23-30). In Belgium, the CEFR is integrated into curriculum planning and language evaluation, notably through the *Socles de Compétences* and exams such

[11] The Common European Framework of Reference for Languages (CEFR), developed in 2001 by the Council of Europe, offers a clear and unified guide for language teaching, learning, and assessment across Europe. It encourages plurilingualism, intercultural understanding, and shared values. Based on an action-oriented approach, it sees learners as social agents and supports learner-centered teaching through general and communicative competences (Council of Europe, 2001: 4).

as the CEB[12] and CE1D[13]. It also informs teacher training, promoting coherence and equity in language education (2001: 13).



Screenshot 2: CEFR levels of proficiency (Council of Europe, 2001: 23).

2.1.2 The action-oriented approach

The CEFR promotes an action-oriented approach, which sees learners as social agents who use language to perform real-world tasks in specific contexts, rather than simply acquiring knowledge about the language. However, the only pedagogical model suggested in the official documents is project-based learning, which is often difficult to apply to the structure of a didactic sequence. That is why Mr Simons' course played an important role in helping to meet this action-oriented perspective through the use of textual genres. He emphasizes that textual genres are not only tools for communication but also instruments for action within specific social contexts, fully aligned with the CEFR's expectations (2018: 18). Since the framework views learners as social actors, it is essential to prepare them for authentic communicative practices through genres like CVs, interviews, cover letters, or debates. These formats help students take an active role in society and foster essential citizenship skills (2018: 229). For this reason, textual genres were included as one of the criteria in the evaluation grids I developed for this study.

Simons (2023) also points out that the concept of textual genres was not explicitly mentioned in the legal documents before 2017, and even today, it remains vague. While some Unités d'Acquis d'Apprentissage (UAAs) (see section 2.4.5) do include examples of genres, the term itself is neither clearly defined nor supported by concrete methodological guidance, which

[12] Certificat d'Etude de Base: It is a standardized external exam for all student completing the primary education (WBF, n.d., enseignement.be).
[13] Certificat d'Etude du 1er Degré de l'enseignement secondaire: It is a standardized external exam for all students completing the first stage of secondary education, with identical content and evaluation criteria across subjects (WBF, n.d., enseignement.be).

makes its implementation challenging for teachers (Simons, 2023: 154). This lack of clarity reinforces the importance of integrating explicit instruction on textual genres, as they offer a practical way to align with the CEFR's action-oriented approach. However, it is also important to recognize the limits of this approach in classroom settings. Although it encourages real-world language use, it cannot always be fully applied in school contexts. The framework itself acknowledges this by describing classroom activities as "un faire semblant accepté volontairement", which is a kind of purposeful role-play that allows students to practise meaningful communication in simulated, controlled conditions (WBF, 2017: introduction). While this strategy helps connect classroom learning with real-life language use, it also highlights the practical constraints faced by teachers in everyday practice.

2.1.3 ChatGPT in the Common European Framework of Reference for Languages

The original CEFR document makes no reference to ChatGPT, artificial intelligence, or online learning, as it was developed in a pre-digital context. In the updated 2018 companion volume, there is still no mention of ChatGPT or AI tools. However, the document does include descriptors for "online interactions", acknowledging that communication can take place through digital means (Council of Europe, 2018: 99). Despite this, the pedagogical use of AI or generative tools is not addressed. This highlights the need for further updates to incorporate digital literacy and AI-assisted language learning, ensuring the CEFR evolves in line with current technological realities.

2.3 The place of artificial intelligence and digital tools in the current curricula

As a reminder, the modern languages framework in the French-speaking Community of Belgium defines what must be taught at each stage of compulsory education. It provides a clear structure for language teaching and learning by aligning with the CEFR and emphasizing real-life communication. Learning outcomes are organized according to CEFR levels and Unités d'Acquis d'Apprentissage (UAAs). The 2017 revision of the framework highlights the importance of learning progression, student engagement, and an action-oriented approach, where learners carry out meaningful and realistic tasks.

In parallel, each educational network in the French-speaking Community is responsible for developing its own curriculum. While the framework establishes what students are expected to

learn in terms of competences and knowledge, the curriculum focuses on how these goals should be achieved in the classroom. It provides pedagogical guidance and teaching strategies to help teachers implement the framework effectively.

The following sections will examine whether the study plans include any reference to digital tools or AI designed for teachers, since the curriculum's main function is to guide the implementation of teaching practices.

2.3.1 Occurrences of "ChatGPT", "Intelligence artificielle", "en ligne" and "Internet" in the curricula of the free subsidised network

	'ChatGPT'	**'Intelligence artificielle"**	**'en ligne'**	**'Internet'**
Langues modernes I, II, III. Programme. Humanités générales et technologiques 2e et 3e degrés (2018) (166 pages)	0	0	8	4
Langues modernes. Formation générale commune. Programme. Humanités professionnelles et techniques 2e et 3e degrés (2017) (129 pages)	0	0	6	4

Table 1: Occurences of "ChatGPT", "Intelligence artificielle", "en ligne" and "Internet" in the curricula of the free subsidised network

In both curricula issued by the free subsidised network, the term "en ligne" appears in contexts related to oral production involving interaction or to describe situations where a message may be produced or received online, such as through digital messaging, but without referring to any pedagogical platforms or teaching tools. Similarly, "internet" is mentioned in relation to the digital environment, rather than as a space for educational resources or teacher development. Notably, the documents contain no mention of "ChatGPT" or "artificial intelligence", which highlights a clear absence of integration of AI tools in the pedagogical guidance. While there is some recognition of online communication as part of language use, there is no reference to the potential use of digital platforms by teachers for accessing pre-designed or AI-assisted didactic materials. This suggests that current curricula acknowledge the digital reality of learners'

language practices but have yet to reflect the pedagogical implications of recent technological advances such as generative AI.

2.3.2 Occurrences of “ChatGPT”, “Intelligence artificielle”, “en ligne” and “Internet” in the curricula of the Wallonia-Brussels Education (WBE) network

	‘ChatGPT’	**‘Intelligence artificielle’**	**‘en ligne’**	**‘Internet’**
Langues modernes. Programme d’études. Humanités générrales et technologiques 2e et 3e degrés (2020) (448 pages)	0	0	19	5
Langues modernes. Programme d’études. Humanités professionnelles et techniques 2e et 3e degrés (2020) (386 pages)	0	0	12	9

Table 2: Occurences of “ChatGPT”, “Intelligence artificielle”, “en ligne” and “Internet” in the curricula of the WBE network

In both curricula published by the WBE network, the term “en ligne” is primarily associated with resources such as online videos and online dictionaries, indicating that teachers are encouraged to incorporate internet-based materials into their language teaching. These references reflect a general openness to using web content to enrich pedagogical practices. However, there is no explicit mention of structured online educational platforms or their formal integration into teaching scenarios. In some cases, references to “sites en ligne” include tools like Quizlet and Memrise (WBE, 2020: 49), which are digital tools focused mainly on vocabulary memorization and individual learning, rather than comprehensive, teacher-oriented didactic systems.

The term “internet” is used to reflect the broader technological dimension of language education. For instance, the curricula include guidelines such as: “utiliser ces supports audio(visuels) pour entraîner d’autres compétences : tâches de lecture apparentées (lectures de critiques, de témoignages, de posts sur Internet…)” (WBE, 2020: 130-133), which suggests the potential use of online content to support various language activities. In another example from the general and technological humanities curriculum, a remediation suggestion reads: “tu peux

aussi trouver des documents sonores et leur script sur Internet. Ton professeur pourra te donner une liste de sites intéressants" (WBE, 2020: 439), implying that teachers may guide students in selecting relevant online resources.

Despite this openness to online materials, a key observation is the complete absence of the terms "artificial intelligence" and "ChatGPT" in both curricula. This omission suggests that AI tools have not yet been formally recognized or integrated into the official pedagogical discourse. While teachers are encouraged to draw from the internet for enrichment purposes, there is no indication of the curricular use of generative AI tools, whether for content creation, personalized practice, or support in assessment. Given the increasing use of such tools in educational contexts, this absence highlights a potential gap between current teaching practices and emerging digital innovations, which may require attention in future curriculum updates.

2.4 Core pedagogical features of the Wallonia-Brussels Federation framework

2.4.1 The Wallonia-Brussels Federation triplet: "Connaître-Appliquer-Transférer"

The "Connaître, Appliquer, Transférer" triplet is a model used across disciplines, not only in language education. This learning progression begins with the acquisition of essential knowledge and the development of their meta cognitive skills (Connaître), followed by the practical application of this knowledge in routine situations (Appliquer), and culminates in the ability to use it flexibly in unfamiliar or complex contexts (Transférer). The goal of this model is to ensure that students not only retain information, but also develop the ability to use it critically and independently, transferring what they have learned to new domains or problems (WBF, 2017: Avant-propos).

2.4.2 Sociocultural, citizenship and digital dimensions in education

In addition to the cognitive stages, the framework also includes several important educational dimensions that support comprehensive language learning. The sociocultural dimension helps students become more aware of cultural diversity, encouraging them to question stereotypes and recognize key cultural references for effective communication. The citizenship dimension invites learners to reflect on topics like media literacy, sustainability, and environmental responsibility, helping them become informed and active citizens. Finally, the digital dimension

focuses on developing digital skills, encouraging students to use technology in a critical and constructive way in their language learning (WBF, 2017: Avant-propos).

2.4.3 Target levels

Moreover, the framework clearly outlines target levels for language proficiency, in alignment with the CEFR. These levels act as indicators for what learners should achieve at various points in their schooling, helping to guide teaching and assessment in a coherent and structured manner (WBF, 2017: Avant-propos).

	Transition							
	LM1 4H		LM2 4H				LM3 4H	
	2e degré	3e degré	2e degré		3e degré		3e degré	
			GERMANIQUES	ROMANES	GERMANIQUES	ROMANES	GERMANIQUES	ROMANES
CA	B1 (-)	B1 (+)	A2 (+)	A2 (+)	B1 (-)	B1 (+)	A2 (+)	B1 (-)
CL	B1 (-)	B2 (-)	A2 (+)	A2 (+)	B1 (-)	B2 (-)	A2 (+)	B1 (-)
EE	B1 (-)	B1 (+)	A2 (+)	A2 (+)	B1 (-)	B1 (+)	A2 (+)	A2 (+)
EOEI	B1 (-)	B1 (+)	A2 (+)	A2 (+)	B1 (-)	B1 (-)	A2 (+)	A2 (+)
EOSI	B1 (-)	B1 (+)	A2 (+)	A2 (+)	B1 (-)	B1 (-)	A2 (+)	A2 (+)

Screenshot 2: Levels expected at the different stages of transitional education (WBF, 2017: avant-propos).

	2H		4H		Formation commune 2H	
	2e degré	3e degré	2e degré	3e degré	2e degré	5e - 6e - 7e
CA	A2 (-)	A2 (+)	A2 (+)	B1 (-)	A1 (+)	A2 (-)
CL	A2 (-)	A2 (+)	A2 (+)	B1 (-)	A1 (+)	A2 (-)
EE	A2 (-)	A2 (+)	A2 (-) *	A2 (+) *	A1 (+)	A2 (-)
EOEI	A2 (-)	A2 (+)	A2 (-) *	A2 (+) *	A1 (+)	A2 (-)
EOSI	A2 (-)	A2 (+)	A2 (-) *	A2 (+) *	A1 (+)	A2 (-)

Screenshot 3: Levels expected at the different stages of technical and professional education (WBF, 2017: avant-propos).

The framework for modern languages identifies five key competences to be developed in a balanced way: listening, reading, oral production with and without interaction, and writing. These skills are meant to be practiced regularly to ensure language development (WBF, 2017: Avant-propos).

2.4.4 Strategies

The "Référentiel Langues Modernes 2017" highlights the essential role of strategies in language learning, presenting them as core components that enable learners to actively mobilize their

knowledge and competences in communicative situations. Far from being instinctive or automatic, strategies are described as organized, conscious, and goal-directed operations that learners use to plan, manage, and evaluate their communication (WBF, 2017: Introduction).

The framework clearly states that strategies need to be taught explicitly. Students are not expected to develop them on their own. Instead, teachers should model, guide, and practice these strategies with their students. This helps learners gradually understand and use them independently when they face communication challenges (WBF, 2017: Introduction).

Moreover, strategies are closely linked to learner autonomy and metacognitive awareness. Being able to plan, monitor, and reflect on one's work is seen as a key sign of autonomy. By teaching strategic behaviour, teachers help students not just complete tasks but also understand and manage their own learning. This idea fits well with the action-oriented approach of the CEFR, which sees language learning as more than just knowing rules: it is about becoming an active and reflective language user (WBF, 2017: Introduction).

The curriculum also highlights that many strategies are transversal, meaning they go beyond language learning and can be used in other subjects. Strategies like planning, reformulating, dealing with missing information, and evaluating results are useful in both school and life situations. Because they apply in many contexts, these strategies are especially important today, as students need to handle complex and changing communication environments (WBF, 2017: Introduction).

Finally, the document organizes strategies according to the different phases of communication:

- Before communication (planning and anticipating needs),
- During communication (adapting, compensating, managing misunderstandings),
- After communication (evaluating and correcting).

This cyclical view of communication is crucial because it shows that using strategies is an ongoing and flexible process (WBF, 2017: Introduction).

In the context of digital and AI learning, such as the use of ChatGPT, the importance of strategic competence is also pronounced. Interacting effectively with AI tools requires learners to plan their inputs, critically assess the relevance of outputs, and adjust their queries or understanding accordingly. Thus, explicitly teaching strategies remains fundamental to preparing students for responsible and critical use of emerging technologies in education.

2.4.5 Unités d'Acquis d'Apprentissage (UAAs)

The document also emphasises the importance of transversal strategies, which are general learning methods that apply across different subjects. These strategies, whether used within a single discipline or across various fields, must be explicitly taught and practiced, as students do not naturally acquire them on their own. Each UAA is structured to support this progression, identifying and mobilizing the knowledge, skills, attitudes, and strategies needed throughout the learning process. This approach helps students develop not only language competences but also transferable skills relevant across disciplines and real situations (WBF, 2017: Avant-propos).

2.5 Conclusion

The elements outlined above — the CEFR, the Décret Missions (1997), the modern languages framework and curricula, as well as the pedagogical principles such as the spiral learning approach, learning strategies, and the Unités d'Acquis d'Apprentissage (UAAs) — form the legal and educational foundation upon which language education is structured in the French-speaking Community of Belgium.

These frameworks explain not only what students should learn, but also how teaching and assessment should be planned to help them become communicative, independent, and flexible language users. They highlight the importance of action-based learning, teaching strategies clearly, and including sociocultural, citizenship, and digital aspects.

Because of their importance, these elements were included in the evaluation grids used in this dissertation. They act as reference points to judge the quality and relevance of the didactic sequences created by ChatGPT. This helps ensure that the evaluation stays connected to official standards and is useful for real teaching in the Belgian system.

2.6 The Presentation-Fixation-Exploitation (PFE) model

Since the analysis of the didactic sequences produced by ChatGPT is based on the PFE model, it is essential to first provide a clear explanation of this model and its key characteristics.

2.6.1 Definition and Origins

The Presentation-Fixation-Exploitation (PFE) model is a structured pedagogical approach developed at the University of Liège, primarily by Wautié-Franck (1988) and later refined by Simons (1997, 2015) (Simons, 2023: 29). This framework is designed to structure the teaching of foreign languages in a systematic and progressive manner, ensuring that learners move from comprehension to controlled practice and finally to open production.

2.6.2 Why Focus on the PFE model?

The decision to focus on the PFE approach in this study is motivated by both its theoretical alignment with recognised models and its practical value in teacher education.

First, its resemblance to the Present-Practice-Produce (PPP) model, frequently found in Anglo-Saxon contexts, offers a familiar structure for many language teachers. Like PPP, PFE proposes a gradual sequence that begins with exposure to new language forms (Presentation), continues with controlled practice (Fixation), and ends with freer use and integration (Exploitation) (Simons, 2023: 30).

Second, PFE aligns closely with the Wallonia-Brussels Federation's "Connaître-Appliquer-Transférer" triplet. Each of these frameworks emphasizes progression, autonomy, and transfer, but with subtle differences in focus. For instance, the WBF approach places strong emphasis on metacognitive reflection, both before and after tasks (WBF, 2017: 5), while the PFE model sticks more closely to a structured teaching method.
The convergence of these models strengthens the pedagogical validity of the PFE approach and justifies its use as a reference in educational practice.

Finally, the PFE model is part of the initial teacher training at the University of Liège and is explicitly promoted as a planning tool for language sequences. In a context where the programs grant substantial pedagogical freedom, such structured references are invaluable. They serve as guidelines for new teachers, offering a clear path for planning while remaining adaptable to individual teaching styles.

2.6.3 The structure of the PFE model

a. Presentation

The Presentation Phase is the first stage of the PFE model, and it serves as the foundation for the subsequent stages. This phase is carefully structured into three key steps to ensure that learners receive a clear and comprehensive introduction to the new linguistic material.

- Contextualization: The first step in the Presentation phase involves introducing the linguistic elements in a meaningful and engaging way. The objective here is to create a sense of necessity for the new material, encouraging learners to perceive it as relevant to their communication needs.
- Clarification: The second step focuses on the explicit explanation of linguistic elements, which may include vocabulary, grammar rules, communicative functions, textual genres or meta cognitive strategies. During this step, students may engage in activities such as extensive, selective, or intensive comprehension tasks, depending on the complexity of the material.
- Memorization: The final step in this phase aims to facilitate the retention of newly introduced content. This can be achieved through structured exercises, at-home study, or interactive classroom activities (Simons, 2023: 31).

While this phase resembles the "Present" stage in the PPP model, it differs slightly from the "Connaître" phase in the WBF triplet, which places a stronger emphasis on identifying the resources mobilized for the final task and, by the end of the sequence, on the metacognitive skills that contributed to its successful completion (WBF, 2017: 5).

b. Fixation

The Fixation Phase serves as an intermediary stage between the initial introduction of new material, its clarification and its application in real communicative contexts. The purpose of this phase is to allow learners to engage in structured and progressive practice, reinforcing their understanding and internalizing the linguistic elements before they are required to use them more spontaneously.

During this phase, students complete a variety of exercises, ranging from highly controlled activities to semi-open tasks. These exercises target lexical, grammatical, functional, strategic or genre-specific aspects of language, ensuring that students gain confidence and fluency in their use of the new material. The gradual nature of this stage allows learners to build their skills

progressively, moving from simple manipulations of language structures to more flexible exercises that encourage creative language use (Simons, 2023: 31).

This phase aligns with both the “Practice” stage of the PPP model and the “Appliquer” phase in the WBF framework. In both models, learners engage with structured activities that promote repeated use of new language forms in relatively predictable contexts, a necessary step before transferring this knowledge to more open tasks (WBF, 2017: 5-6).

c. Exploitation

The final phase of the PFE approach is crucial for ensuring that learners move beyond controlled exercises and begin to apply their knowledge in authentic communicative situations. This phase is not solely about assessment; rather, it aims to foster an integrative approach to language learning, allowing students to demonstrate their ability to use the language effectively in new contexts.

This phase is guided by three fundamental principles:

I. Moving beyond the initial document or context in which the linguistic elements were introduced. This ensures that learners are not simply reproducing memorized content but are instead adapting their knowledge to different situations.
II. Encouraging autonomous production, which involves engaging students in oral or written tasks that require them to actively use the language they have acquired.
III. Reducing teacher control to maximize learner autonomy (Simons, 2023: 32).

By integrating these principles, the Exploitation Phase ensures that learners actively use the language, rather than just receive information, and adapt their skills to different communication needs.

This mirrors the “Production” stage in PPP, where students practice speaking freely, and the “Transférer” phase in the WBF approach, which includes real-life tasks that require deeper thinking (WBF, 2017: 7).

d. Evaluation and Remediation phases

After completing the third phase of the sequence, which should serve as a diagnostic tool to determine whether students have achieved the intended learning outcomes. It is essential to include extra activities to help students who struggled and support them in reaching the expected level. After that remediation, an evaluation should be organised. This certificative task

targets a competence, as well as the vocabulary, language functions, textual genre and/or grammar developed during the sequence (Simons, 2023: 90-91).

2.6.4 Adaptability and relevance

One of the key strengths of the PFE model lies in its flexibility. While its structure is linear (P-F-E), it allows for adapted sequences such as P-P-F-E or even P-F-P-F-E-E, depending on the lesson duration, learning objectives, or the complexity of the materials used. Nevertheless, a fundamental principle must be upheld to ensure pedagogical coherence: the progression should move from presentation to fixation and then to exploitation (Simons, 2023: 32–37).

This adaptability makes it particularly valuable in both face-to-face and online contexts, where teachers often need to adjust their planning on the fly. Compared to less predictable models (such as the problem-based model) PFE offers a clear structure that facilitates anticipation and organization, especially for novice teachers or in structured environments like digital platforms.

The problem-solving framework is based on an inductive and learner-centered approach that begins with the final communicative task and works backwards to identify required knowledge and skills. This dynamic process makes it difficult to pre-plan AI prompts or responses, especially when learners' needs vary or are hard to diagnose. Furthermore, the critical "état des ressources" phase relies on learners' ability to reflect on and articulate prior knowledge, which is something ChatGPT cannot easily interpret or support (Simons, 2023: 39-44).

In addition, teachers can adapt the phases based on several factors, including: the available classroom time, the complexity of the material, the learner's prior knowledge, and the level of interaction required in comprehension activities, which may vary depending on whether students are working with written texts, audio recordings or interactive dialogues (Simons, 2023: 32).

This adaptability makes PFE a dynamic teaching framework, capable of being customized to meet the diverse needs of learners in different educational settings.

2.6.5 Conclusion

The PFE model offers a clear, step-by-step, and flexible way to teach languages. It shares similarities with the PPP model and the WBF approach. Its focus on progression, autonomy, and transfer aligns it with competency-based education.

By giving teachers a method they can adapt to different situations, the PFE approach supports their independence and helps students move smoothly from understand to speaking.

3. Analysis of a didactic sequence created by ChatGPT

RQ6: ***Is ChatGPT capable of creating a coherent didactic sequence, adapted to language education in WBF, when guided by a language teacher?***

3.1 Methodology

This study aims to evaluate whether ChatGPT is capable of producing a coherent didactic sequence that meets the legal and curricular expectations of language education in the French-speaking Community of Belgium. To do so, a systematic approach was adopted, combining careful prompt design, controlled content generation, and structured evaluation.

Before asking ChatGPT to generate any content, it was necessary to define the conditions under which the tool would be tested. Since the goal of the study is to assess ChatGPT's ability to create sequences that reflect real classroom practices, the test was designed to mirror a typical situation faced by teachers: having a topic in mind and needing to build a complete sequence from it. In this case, the selected theme was "Gap Year" and ChatGPT was asked to verify whether the topic aligned with the learners' CEFR level, and then to produce a twelve-lesson sequence.

The first phase consisted of preparing the model for the task by constructing precise and detailed prompts. Before asking ChatGPT to produce any didactic material, it was explicitly provided with explanations of key pedagogical frameworks and official documents relevant to the Belgian educational context. These included the PFE framework commonly used in lesson planning and taught in the teacher training program at the University of Liège, which is

described in the section 3.6. The structure and descriptors of the CEFR, and the concept of UAAs as outlined in the "Référentiel Langues Modernes 2017", were also informed to the tool. This preliminary phase ensured that ChatGPT had access to all essential pedagogical and legal references to guide the production of a sequence consistent with current educational standards.

Once the foundational information was introduced, ChatGPT was tasked with designing a complete didactic sequence. The request specified that the sequence should target a fifth-grade class in Liège, where English is taught as the first foreign language. The theme selected for the sequence was "Gap Years", with a grammatical focus on the different forms used to express the future in English. ChatGPT was asked to plan a twelve-lesson sequence. It first generated a general overview of the sequence, describing the overall progression and main objectives. Following this, each individual lesson was created separately, detailing the learning goals, activities, targeted skills, and expected outcomes for every session.

Following the production of the sequence, the third phase of the study involved an evaluation of ChatGPT's work. To do so, evaluation grids specifically developed for this research were applied. These grids were constructed based on the legal expectations and the PFE model. It focuses on several aspects such as the overall coherence of the sequence, the alignment with official standards (for instance, the CEFR levels), the quality of the activities etc. The different grids and their descriptions are explained in section 4.3.

Through this procedure, the study seeks to determine the extent to which ChatGPT can autonomously generate didactic sequences that not only demonstrate internal coherence but also respect the pedagogical and legal requirements established for language education in the French-speaking Community of Belgium.

3.2 Pre-test phase and prompt adjustments

Before the main test was launched, a pre-test phase was conducted to refine and adapt the prompts for greater clarity and effectiveness. In the first attempts, ChatGPT often created lessons with only one large activity and little variety. To fix this, the prompts were revised to specify that a standard lesson lasts approximately 50 minutes but should be realistically planned over 45 minutes to account for classroom management and transitions. Additionally, the tool was explicitly informed of the legal and pedagogical expectations specific to the Belgian context. This included explanations of the PFE model, CEFR level descriptors, and the role of

UAAs in structuring and evaluating language learning. This preparatory work was essential to guide ChatGPT in generating a didactic sequence that was not only logically organized but also aligned with real teaching requirements.

3.3 Evaluation grids

The following evaluation grids were created to assess whether ChatGPT is capable of generating a coherent didactic sequence based on the PFE model. Their purpose is to evaluate if ChatGPT-generated sequences follow a logical progression, respect key didactic principles, and align with professional teaching standards.

Each grid focuses on a specific aspect of the sequence, ensuring that all phases are correctly structured and effectively implemented. Moreover, a general evaluation grid verifies the overall quality of the sequence, including its alignment with official curricula and teaching methodologies. It is essential to note that no grid was developed for an evaluation, as no evaluation task will be created or analysed with the scope of this study.

By using these grids, we can systematically test ChatGPT's capabilities and determine whether it can serve as a reliable tool for lesson planning in language education in French-speaking Belgium.

3.3.1 Evaluation grid 1: General grid

Criteria	Indicators	Evaluation	Points	Remarks
Presence of the three phases	The didactic sequence explicitly includes all three phases as outlined in the PFE framework.	Strongly disagree (0) Disagree (1) / Agree (2) / Strongly Agree (3)	/3	
Logical coherence between the three phases	The sequence shows a clear and progressive link between the phases, with each phase building on the previous one.	Strongly disagree (0) / Disagree (1) / Agree (2) / Strongly Agree (3)	/3	
Clarity of learning objectives	The sequence explicitly states what students will learn, and these objectives are aligned with the activities.	Strongly disagree (0) / Disagree (1) / Agree (2) / Strongly Agree (3)	/3	
Alignment with the CEFR levels	The sequence aligns with the levels of competency mentioned in the CEFR.	Strongly disagree (0) / Disagree (1) / Agree (2) / Strongly Agree (3)	/3	
Alignment with the topic units in the WBF framework	The sequence aligns with the topic units mentioned in the WBF framework.	Strongly disagree (0) / Disagree (1) / Agree (2) / Strongly Agree (3)	/3	
Alignment with the grammar points in the WBF framework	The sequence aligns with the grammar points mentioned in the WBF framework.	Strongly disagree (0) / Disagree (1) / Agree (2) / Strongly Agree (3)	/3	

Presence of the UAA(s) worked	The sequence mentions what "UAA" will be worked throughout the sequence.	Strongly disagree (0) / Disagree (1) / Agree (2) / Strongly Agree (3)	/3	
Quality of learning materials	The materials used are engaging.	Strongly disagree (0) / Disagree (1) / Agree (2) / Strongly Agree (3)	/3	
Use of a communicative approach	The sequence prioritises meaningful communication (prioritising the action-oriented approach).	Strongly disagree (0) / Disagree (1) / Agree (2) / Strongly Agree (3)	/3	
Learning strategies	Elements of metacognition working on receptive strategies (listening/reading) are integrated.	Strongly disagree (0) / Disagree (1) / Agree (2) / Strongly Agree (3)	/3	
	Elements of metacognition working on productive strategies (writing/speaking with or without interaction) are integrated.	Strongly disagree (0) / Disagree (1) / Agree (2) / Strongly Agree (3)	/3	

Table 3: Evaluation grid 1: General grid.

The general grid was created to assess not only the presence of the three PFE phases but also the overall quality and coherence of the didactic sequence. While the PFE model provides a structured teaching model, a sequence must also align with broader pedagogical principles and official requirements in the Wallonia-Brussels Federation.

First, the grid ensures that all three phases (Presentation, Fixation, Exploitation) are clearly present and follow a logical progression. A well-structured sequence must introduce new content, reinforce it through controlled exercises, and allow students to use it autonomously. Without this coherence, learning remains fragmented.

Second, the grid evaluates whether the sequence aligns with official language programs. Since these programs offer pedagogical freedom, as explained in the introduction, the grid verifies if the sequence respects competency-based teaching, particularly the "Connaître-Appliquer-Transférer" approach, which emphasizes knowledge acquisition, practice, and transfer. Its alignment with the CEFR is also evaluated, especially the level of competency, the topic units and the grammar points.

Moreover, the quality of teaching materials and linguistic accuracy are crucial to ensure meaningful and correct language exposure. The grid also assesses whether the sequence follows a communicative approach, encourages student autonomy, and adapts to different learning needs.

This evaluation grid uses structural, teaching, and language criteria to make sure a lesson plan is well-designed and meets school expectations.

3.3.2 Evaluation grid 2: Presentation Phase

Criteria	Indicators	Evaluation	Points	Remarks
Introduction: arouses student interest	The introduction is engaging and relevant to students' experiences.	Strongly disagree (0) / Disagree (1) / Agree (2) / Strongly Agree (3)	/3	
Introduction: creates a need for learning	The activity naturally leads to the necessity of acquiring new linguistic elements.	Strongly disagree (0) / Disagree (1) / Agree (2) / Strongly Agree (3)	/3	
Introduction: facilitates comprehension	The activity includes elements (e.g., vocabulary, reactivation, visual supports) that aid in comprehension.	Strongly disagree (0) / Disagree (1) / Agree (2) / Strongly Agree (3)	/3	
Clarification: uses extensive or selective comprehension strategies	The sequence begins with a general comprehension of the document before moving on to more detailed analysis.	Strongly disagree (0) / Disagree (1) / Agree (2) / Strongly Agree (3)	/3	
Clarification: clearly explains new linguistic elements	Vocabulary is explained explicitly.	Strongly disagree (0) / Disagree (1) / Agree (2) / Strongly Agree (3)	/3	
	Grammar is explained by the inductive approach.	Strongly disagree (0) / Disagree (1) / Agree (2) / Strongly Agree (3)	/3	
	The textual genre is explained explicitly.	Strongly disagree (0) / Disagree (1) / Agree (2) / Strongly Agree (3)	/3	
	The functions are explained explicitly.	Strongly disagree (0) / Disagree (1) / Agree (2) / Strongly Agree (3)	/3	
Memorization : reinforces linguistic retention	The sequence includes activities that encourage students to retain new information.	Strongly disagree (0) / Disagree (1) / Agree (2) / Strongly Agree (3)	/3	

Logical and coherent sequencing	The three sub-phases (introduction-clarification-memorization) are clearly structured and sequentially ordered.	Strongly disagree (0) / Disagree (1) / Agree (2) / Strongly Agree (3)	/3	

Table 4: Evaluation grid 2: Presentation phase

The evaluation grid for the Presentation Phase was developed to ensure that this initial stage of the PFE model effectively introduces new linguistic content in a structured and engaging manner. Since the presentation phase is the starting point of learning, it needs to be clear and organised.

This grid evaluates whether the introduction successfully arouses student interest and creates a need for learning. If students do not see the relevance of the new content, their engagement and retention may be compromised. The clarification phase is also assessed to verify whether new linguistic elements (grammar, vocabulary, functions, textual genre) are explained explicitly and progressively.

Finally, the memorization phase must provide reinforcement activities and clear study strategies to help students internalize the new content. A well-designed presentation phase not only delivers information but also prepares learners for the fixation phase, ensuring that they can actively engage with the language rather than passively receive it.

By assessing these key components, this grid helps determine whether the presentation phase is engaging and structured in supporting student learning.

3.3.3 Evaluation grid 3: Fixation Phase

Criteria	Indicators	Evaluation	Points	Remarks
Progression of complexity of the exercises	The exercises follow a progression from structured drills to semi-open activities.	Strongly disagree (0) / Disagree (1) / Agree (2) / Strongly Agree (3)	/3	
Fixation includes reinforcement	The sequence contains exercises aimed at reinforcing vocabulary/ grammar/ functions/ textual genre learned in the presentation phase.	Strongly disagree (0) / Disagree (1) / Agree (2) / Strongly Agree (3)	/3	
Exercise variety	The sequence includes diverse activities to maintain student motivation.	Strongly disagree (0) / Disagree (1) / Agree (2) / Strongly Agree (3)	/3	
Gradually reduces control over exercises	The transition from structured practice to freer exercises ensures a gradual release of responsibility towards students.	Strongly disagree (0) / Disagree (1) / Agree (2) / Strongly Agree (3)	/3	

Table 5: Evaluation grid 3: Fixation phase

The evaluation grid for the Fixation phase was designed to make sure students get structured, step-by-step practice after being introduced to new linguistic elements. In the PFE model, this phase is essential for reinforcing learning, as it bridges the gap between presentation and autonomous use.

This grid evaluates whether exercises are varied and progressive, moving from controlled drills to semi-open activities that encourage increasing student autonomy. It also verifies whether all key linguistic aspects (vocabulary, grammar, functions, textual genres) are effectively practiced, making sure that students internalize the material before moving to open production.

By assessing these criteria, the grid ensures that the fixation phase is structured, interactive, and prepares students for independent language use.

3.3.4 Evaluation grid 4: Exploitation phase

Criteria	Indicators	Evaluation	Points	Remarks
Context and task plausibility	The proposed task is realistic, clearly situated in a believable context.	Strongly disagree (0) / Disagree (1) / Agree (2) / Strongly Agree (3)	/3	
Encourages transfer and integration of knowledge	The task requires students to reuse all the linguistic elements learned and fixed in new contexts.	Strongly disagree (0) / Disagree (1) / Agree (2) / Strongly Agree (3)	/3	
Moves beyond the initial document and context	The task encourages students to use language elements outside of the specific situation in which they were introduced.	Strongly disagree (0) / Disagree (1) / Agree (2) / Strongly Agree (3)	/3	
Activities reinforce communicative competence	The task focuses on meaningful communication rather than isolated grammar or vocabulary drills.	Strongly disagree (0) / Disagree (1) / Agree (2) / Strongly Agree (3)	/3	
Promotes active linguistic production	The sequence includes oral and/or written tasks that require students to generate messages independently rather than repeating pre-learned content.	Strongly disagree (0) / Disagree (1) / Agree (2) / Strongly Agree (3)	/3	
Increases student autonomy by reducing teacher control	The task is designed to be more student-led, allowing for creative and personalised language use.	Strongly disagree (0) / Disagree (1) / Agree (2) / Strongly Agree (3)	/3	
No single correct response is expected	The task allows for multiple possible answers and interpretations.	Strongly disagree (0) / Disagree (1) / Agree (2) / Strongly Agree (3)	/3	
The quality of exploitation depends on the effectiveness of the previous phases	There is a clear progression from presentation and fixation to exploitation, ensuring students are adequately prepared.	Strongly disagree (0) / Disagree (1) / Agree (2) / Strongly Agree (3)	/3	

Table 6: Evaluation grid 4: Exploitation phase

The evaluation grid for the Exploitation phase was developed to ensure that students transfer and integrate their newly acquired knowledge into authentic communication. This phase is crucial for moving beyond controlled exercises toward autonomous and creative language use.

This grid evaluates whether students are required to reuse linguistic elements in new contexts, ensuring that learning extends beyond the classroom. It also checks if activities gradually reduce teacher control, fostering independent production and encouraging students to express original ideas without rigid structures.

Additionally, it also checks whether tasks help develop real communication skills by focusing on meaningful exchanges instead of just repetitive exercises. A good exploitation phase should accept different correct answers, just like real-life communication, where language is flexible and varied.

By applying this grid, we can determine whether the exploitation phase effectively prepares students for real-world language use.

3.4 Results of the sequence produced by the free version of ChatGPT[14]

To evaluate the didactic sequence produced by ChatGPT, the evaluation grids presented earlier were applied. ChatGPT was instructed to design a sequence for fifth-grade students in a school in Liège, where English is taught as the first foreign language. The topic chosen was “Gap Year”, and the sequence was organized across twelve lessons. The ChatGPT model used was the 4o mini, which offers an unlimited access for free users.

Overall, the sequence achieved a score of 39 out of 99[15], indicating that ChatGPT is relatively effective and relevant in producing pedagogically appropriate content.

In the general evaluation grid, ChatGPT obtained a score of 20.5 out of 33. Although the sequence maintained a degree of overall quality, major issues were observed. Notably, there was a lack of coherence between the phases of the PFE model. ChatGPT proposed a variety of exploitation tasks, such as a debate, a writing task (email), and multiple oral productions, without establishing a clear and consistent final objective. This fragmented approach risks

[14] See Appendix 3
[15] See Appendix 1

creating confusion and frustration among students, as the central learning goal of the sequence was not obvious. In addition, the tool did not provide any concrete listening comprehension activities, making it impossible to assess the quality of authentic materials. As in the premium version, no work on metacognitive strategies for receptive skills was integrated.
The presentation phase was particularly problematic, resulting in a score of 5 out of 30. No clear introduction was provided, no listening comprehension link was offered, and the grammatical content was not taught through an inductive approach. Furthermore, communicative functions were not identified or explained. The sequencing lacked logic and internal coherence, undermining the effectiveness of the presentation stage.

In the fixation phase, the sequence scored 6.5 out of 12. Here, the main weaknesses were the insufficient number of exercises and the lack of a gradual reduction of teacher control. The exercises proposed did not allow for an effective consolidation of the new language structures.

Finally, in the exploitation phase, the sequence obtained a score of 7 out of 24. The weak foundation laid during the earlier phases made it impossible to design a coherent and realistic final task. Moreover, ChatGPT proposed multiple final tasks (debate, writing task, and oral productions) without a clear progression or prioritization. This multiplicity of objectives made the sequence unrealistic and risked creating frustration among learners by blurring the main communicative aim.

3.5 Results for the sequence produced by the premium version of ChatGPT[16]

The same test was conducted with the premium version of ChatGPT (4o), using identical parameters and instructions. The resulting didactic sequence achieved a score of 69.5 out of 99[17], indicating a significantly higher level of quality compared to the version generated by the free model.

In the general evaluation grid, ChatGPT scored 26 out of 33. The sequence respected major legal expectations and demonstrated the use of a communicative, action-oriented approach. However, some weaknesses were observed regarding the coherence between the different

[16] See Appendix 4
[17] See Appendix 2

phases of the PFE model, and no explicit work on metacognitive strategies for receptive skills was included.

In the presentation phase grid, ChatGPT obtained a score of 18.5 out of 30. A key issue arose during the clarification phase: an additional listening comprehension activity was placed incorrectly within the fixation phase, and the corresponding audio resource was difficult to find online. Furthermore, the sequence did not integrate extensive or selective comprehension strategies. While the grammar point (future forms) was introduced inductively, by presenting example sentences and asking students to infer the rules, it was not properly embedded within the listening or reading activities. Some useful expressions, such as linkers, were mentioned, but their communicative functions were not explicitly developed. In addition, the structure of the presentation phase (introduction, clarification, memorization) was not respected. Vocabulary definitions, for example, were given before any listening or reading activity, contrary to the sequence instructions. Moreover, no formal vocabulary definitions were provided despite the initial guidance.

Regarding the fixation phase, ChatGPT scored 6.5 out of 12. The main limitation was the insufficient number and variety of exercises for both vocabulary and grammar reinforcement. Several lessons (notably lessons 8 to 11) were overly focused on preparing the final task and conducting peer reviews, rather than consolidating knowledge through varied and targeted exercises.

In the exploitation phase grid, the sequence achieved a score of 18.5 out of 24. ChatGPT succeeded in respecting important criteria such as the plausibility of the context and task, the transfer and integration of knowledge, the possibility of multiple answers, and the promotion of communicative competence. Nevertheless, student autonomy remained somewhat limited. Although peer reviews were encouraged, the oral production task was heavily prepared in advance, potentially leading to rehearsed or memorized answers, rather than fostering spontaneous active language production.

3.6 Comparison between both versions

The results show clear differences between the premium and free versions of ChatGPT. The premium version scored much higher (69.5/99)[18] than the free version (39/99)[19], producing a more coherent and structured didactic sequence. While both versions had some weaknesses, especially in integrating metacognitive strategies, the premium version showed a better respect for the sequence's progression and created a more realistic and achievable final task. It also allowed for the downloading of generated documents, which made it easier for teachers to access, adapt, and reuse the proposed materials.

In comparison, the free version showed more significant limitations. It lacked clear goals, logical organization, and realistic task progression, which reduced the overall quality of the sequence. Although technically able to suggest external links, when the sequence was first generated, it did not provide any real listening comprehensions or concrete resources, only imagined comprehension questions. It was only after activating the "search" option during the conversation that ChatGPT could offer links to some well-known websites. However, these links were suggested separately and were not fully integrated into the structure of the sequence itself. In the end, while both versions still require human supervision and adaptation, the premium version proved to be much more reliable for producing coherent and pedagogically usable content.

3.7 ChatGPT's didactic sequence without any guidance

3.7.1 Introduction

To explore how ChatGPT performs without any prior guidance, I ran an additional test where I only provided three basic pieces of information to ChatGPT's 4o-mini model: the topic of the sequence ("Gap year"), the grammar point to focus on (future forms), and the students' grade level (fifth year of secondary school). Just like in the tests conducted with both the premium and free versions, I asked the tool to design a complete didactic sequence consisting of twelve lessons. However, this time, I did not explain any legal or pedagogical expectations, nor did I mention specific frameworks like the PFE.

[18] See Appendix 4
[19] See Appendix 3

At first, I carried out this test using my personal ChatGPT account. Although I explicitly instructed the tool to disregard all previous conversations, the sequence it generated strongly reflected elements we had discussed in earlier exchanges. To avoid this potential bias, I repeated the test using a different account, free of any memory history, to ensure that the tool's response would be based solely on the new prompt. Since the sequence was not designed around the PFE model, it was not evaluated using the grids created for this dissertation.

The following section presented the results of both tests, offering insights into how ChatGPT performs when left to generate content independently, with and without memory influence.

3.7.2 Results from the test on a new account (without memory)[20]

Despite the absence of detailed prompts, the sequence generated by ChatGPT demonstrated a number of strengths. All the language skills—listening, reading, writing, speaking with and without interaction—were integrated in a balanced way. Metacognitive tasks were also included, asking students to reflect on what they already knew, what they wanted to learn, and what they had achieved, which aligns with best practices in promoting learner autonomy.

The structure loosely mirrored the PFE model. However, the internal progression of the sequence lacked consistency, as fixation activities occasionally came after exploitation tasks, which disrupted the logic of the sequence. Moreover, the sequence introduced multiple final tasks (a group project, an informal email, and a debate) without clear guidance and support. Indeed, this seems impossible to do in twelve lessons because the task has to be introduced and fixed first. This lack of focus weakened the overall coherence and made the main communicative goal unclear. For instance, the debate was introduced without prior language preparation, making it challenging to implement effectively.

Grammar instruction was handled using an inductive approach, which is recognised as effective in language teaching. Students first encountered the future forms "will" and "going to" in a reading passage, then moved on to a synthesis and exercises. Later, the present continuous was introduced in a second text and consolidated through new activities. The fixation phase showed a progression from closed to open tasks, fostering increasing learner autonomy. However,

[20] See Appendix 5

vocabulary was presented before the listening and reading tasks, which could make it harder for students to understand the context**.**

Listening comprehension was addressed in two proposed activities. One included a YouTube link, but the video was unavailable. The second offered a fictional audio scenario without an actual file. This limitation had already been observed in the test conducted with the free version (with guidance). Reading comprehension involved short thematic texts accompanied by questions. These texts also served as the basis for grammar work.

The final group project was relatively well-developed, including peer feedback and structured preparation over two to three lessons. Still, because the textual genre had not been clearly introduced and guided from the beginning, its role in the sequence remained ambiguous.

3.7.3 Results from the test on my account (with memory)[21]

When running the same test using my personal account (where ChatGPT had memory active), the generated sequence was surprisingly aligned with the legal and pedagogical expectations of the French-speaking Community of Belgium. Despite having been explicitly instructed to "forget" previous exchanges, the output included all three phases of the PFE model, integrated metacognitive strategies, respected CEFR levels, and incorporated the strategic teaching of grammar through inductive methods.

The sequence began with a clear and engaging introduction, featuring activities such as brainstorming and quizzes to activate prior knowledge and introduce the topic. However, this introductory phase extended over an entire lesson, which may be excessive for typical classroom constraints. While no full reading comprehension text was provided, ChatGPT suggested themes like "Volunteering in Costa Rica" and proposed related comprehension questions that followed a logical progression: from open-ended understanding to more detailed analysis.

For the grammar point, the sequence applied an inductive approach: students were presented with a short paragraph and asked to identify the grammatical structures, specifically the future forms. This discovery stage was followed by a synthesis of grammar rules and corresponding exercises, showing a coherent pedagogical flow. A listening comprehension activity was also

[21] See Appendix 6

included. However, instead of a real audio file or a working link, ChatGPT created a fictional listening titled “Planning my gap year: interview with Sarah” and even provided an excerpt to analyze grammar points in context. This illustrates the tool’s potential to simulate realistic materials but also exposes a limitation in its inability to embed authentic resources without direct access to external platforms.

Interactive and communicative activities, such as a Jigsaw, were included to consolidate grammar and foster oral communication. The fixation phase was well-structured, offering a progression from closed to more open-ended exercises, supporting gradual learner autonomy. Writing tasks were scaffolded with planning support, time for drafting, and opportunities for peer feedback over two to three lessons. The final task required students to create a brochure in groups and present it to the class, with three lessons dedicated to its preparation. Importantly, the textual genre was introduced and practiced with the students beforehand, ensuring they understood its structure and purpose. Throughout the sequence, self- and peer-assessment activities were encouraged, aligning with current educational recommendations that promote learner autonomy and reflection.

Moreover, the sequence covered all four communicative skills targeted by the official framework: listening and reading comprehension, writing, and oral production both with and without interaction. Despite the absence of explicit pedagogical guidance in the prompt, ChatGPT managed to generate a sequence that was coherent, balanced, and aligned with several key principles of language teaching.

What stood out most in this test, however, was the fact that many of the issues identified in earlier evaluations (such as the lack of fixation exercises, insufficient guidance, and absence of an inductive grammar approach) had been addressed. This occurred even though no specific instructions were given in the prompt, which raises questions about whether the tool was influenced by previous exchanges. It seems that some of the feedback and legal or pedagogical input provided during earlier interactions were retained or echoed in the output, despite the instruction to ignore memory. This strongly suggests that memory may still have played a role, either through direct recall or through the model’s capacity to draw from long-term conversational patterns within the same account.

3.7.4 Reflection and limitations

The comparison between the two tests raises important questions about ChatGPT's internal functioning and the influence of memory on its responses. Both sequences demonstrated several strong features, including the integration of all five communicative competences, the use of metacognitive reflection activities, and a generally coherent inductive approach to grammar. These elements align well with core principles of language education, despite not being explicitly prompted.

However, the sequence generated on my main account (where memory was active)[22] stood out in particular. Although ChatGPT was explicitly asked to ignore any prior exchanges, the tool produced a sequence that not only followed the general structure of the PFE model but also addressed several issues previously highlighted in the earlier tests. It avoided overloading the exploitation phase, improved the quantity and progression of fixation exercises, respected the inductive method of grammar teaching, and included clearer guidance. These improvements suggest that the tool had retained and incorporated my earlier feedback, even when told to "forget" previous conversations. In other words, the memory function appeared to override my request for a fresh start, reflecting ChatGPT's capacity to adapt based on past interactions.

This behavior can be seen as both a strength and a limitation. On the one hand, it shows that ChatGPT has the potential to evolve into a more responsive and tailored teaching assistant when used over time by the same user. On the other hand, it raises concerns about transparency and control over what the tool remembers and how it influences future outputs, especially in research contexts where neutrality is essential.

In contrast, the sequence created using a different account, without memory[23], also showed some internal coherence and respect for key teaching principles, though it presented more inconsistencies. For instance, the multiple final tasks, the reversed sequencing of fixation and exploitation phases, and the lack of authentic resources (particularly for listening comprehension) highlight its limits when working without detailed prompts or accumulated feedback.

[22] See Appendix 6
[23] See Appendix 5

These results suggest that while ChatGPT can produce valuable didactic content even without explicit guidance, its true potential becomes more evident when it builds on prior interactions. This dual capacity reinforces the idea that ChatGPT is not a fully autonomous teaching tool but rather a collaborative assistant that performs best when guided and reviewed by an informed user.

3.8 Conclusion

The comparative analysis of the didactic sequences produced by ChatGPT under various conditions (free version, premium version, and unguided tests with and without memory) highlights clear differences in quality, coherence, and pedagogy. Among these, the sequence generated by the premium version stands out as the most complete and reliable. It not only respected the legal and pedagogical requirements of the French-speaking Community of Belgium but also demonstrated a structured progression, clear communicative goals, and thoughtful integration of all five language competences. Even with a few flaws, like missing real audio and uneven use of learning strategies, the sequence remained balanced and effective, showing the strengths of the premium model.

The test conducted with no guidance on the user's main account (where ChatGPT's memory was active) also produced a surprisingly strong sequence. Although no pedagogical instructions were given, the output aligned closely with key educational principles and addressed several issues raised in previous tests. This suggests that the tool retained and applied earlier feedback, resulting in a more refined and coherent sequence. While not superior to the premium version in overall structure, this result highlights the adaptive potential of ChatGPT when used consistently by the same teacher.
In contrast, the free version showed clear limitations. It lacked internal coherence, failed to provide a clear communicative goal, and offered incomplete or fictional teaching resources. These weaknesses significantly reduced its classroom applicability and overall reliability, highlighting the importance of using more advanced versions for professional contexts.

In summary, these results indicate that ChatGPT is capable of generating quality didactic content, particularly in its premium version or when benefiting from accumulated feedback through memory. However, they also point to the importance of human oversight: even the most advanced version requires the critical eye of a trained teacher to guide, adapt, and

contextualise the material. ChatGPT can be a powerful assistant, but not a substitute for professional judgment and expertise.

4. A survey[24] of (future) language teachers in Liège: uses and perceptions of ChatGPT

4.1 Introduction

This section aims to explore the results of the survey conducted among (future) language teachers in Liège (Wallonia, Belgium), and to answer several of the research questions and sub-questions formulated in this study. These questions focus on understanding how these educators interact with ChatGPT, whether they use it in their teaching practices, and what their opinions are regarding its use.

The survey was designed with two main objectives. First, it aimed to gather general information about the respondents, including their gender, age, years of teaching experience, the languages they teach, and their professional status (whether they are students, teachers, or both). This information helps provide context to better analyze trends in AI usage across different profiles.

The second and more central part of the survey focused on the use of ChatGPT in language teaching. Participants were asked whether they use the tool, and if so, why and how often they use it. For those who do not use ChatGPT, the survey aimed to uncover the reasons behind their reluctance or lack of interest. The questions also examined whether educators modify the responses generated by ChatGPT, why they feel the need to do so, and whether they have access to the free or premium version of the tool. Another key aspect was to assess whether they had received any formal training on how to use AI in education and whether they felt such training would be beneficial.

While this survey provides valuable insights into the perception and use of ChatGPT in language education, it is important to acknowledge its limitations. Since it was conducted exclusively among language teachers and teacher trainees in Liège, its findings may not be fully representative of teachers in other regions of Belgium or beyond. Moreover, as participation was voluntary, the sample may include a higher proportion of respondents who are already

[24] See Appendix 7

interested in AI or educational technology, which could influence the results. In addition, some responses may have been affected by social desirability bias, which is the tendency for participants to give answers they believe are more socially acceptable or professionally valued, rather than reflecting their actual practices or opinions. This bias may be reinforced by the academic context of the study, as respondents knew their answers would be analysed as part of a university dissertation likely to be read by professors.

Despite these limitations, this survey offers a crucial perspective on how AI is perceived in the field of language education and how (future) teachers are adapting (or hesitating to adapt) to the emergence of tools like ChatGPT. The findings will help shed light on AI's role in modern teaching and contribute to the broader discussion on its potential and challenges in the educational landscape.

4.2 Methodology

In order to analyse the usage habits and perceptions of ChatGPT among (future) language teachers in Liège, I decided to conduct an online survey. This method was chosen primarily for its efficiency and accessibility, allowing me to reach a wide audience of teachers and teacher trainees across different institutions. An online format also ensured that respondents could complete the questionnaire at their own pace, making it more convenient for them and increasing the likelihood of obtaining thoughtful and reliable responses. This facilitates the results' analysis and the presentation of them.

4.2.1 Designing the Survey

The development of the survey was a collaborative effort. I worked alongside Lisa Gayet, another student conducting research in the field of language education. Together, we designed the first section of the questionnaire, which aimed to collect general information about the respondents. This part included questions related to gender, age, years of teaching experience, languages taught, and professional status (whether they were students, teachers, or both). This section was essential in providing contextual data that would later allow us to analyse trends based on different profiles. The decision to combine our surveys into a single one was made by Mr Simons to avoid redundancy because similar questionnaires would have been sent to the same target audience, which would have reduced response rates and the overall success of our data collection.

After completing this shared section, we each developed a separate part of the questionnaire according to our respective research focuses. The third and final part, which I created, was specifically dedicated to investigating the use of ChatGPT in language education. By structuring the survey in this way, we ensured that each of our research objectives was addressed while maintaining a coherent and logical progression throughout the questionnaire.

The creation of the survey was guided by expert recommendations. We sought advice from Professor Simons, who suggested that we contact Audrey Renson, a PhD student and assistant in the Department of Didactics of Modern Languages, who taught us how to create a questionnaire. Her insights were particularly valuable in helping us design clear, structured questions to gather useful and accurate information. She also showed us several survey platforms, such as SurveyMonkey, which offers advanced features for collecting and analysing data. However, due to financial constraints, we opted for Microsoft Forms, which, while free, still provided all the necessary tools for conducting a structured and efficient survey. It allowed us to divide the survey into different sections to reflect our separate research focuses, and to include an unlimited number of questions in various formats. Additionally, Microsoft Forms automatically generated graphs for each question and exported the results to an Excel file, which helped us organise and analyse the data in a more structured way.

To allow for a comprehensive analysis, the survey included different types of questions. Some were multiple-choice questions (MCQ) where respondents could select one answer, while others were open-ended questions, allowing participants to elaborate on their thoughts and experiences. In addition, some questions followed a Likert scale format, enabling respondents to express their level of agreement or frequency of use on a graded scale. The Likert scale used in the survey was designed with four levels instead of the traditional five, in order to avoid central tendency bias and encourage respondents to take a clear position.

4.2.2 Pre-Test Phase and Adjustments

Before officially distributing the survey, we conducted a pre-test phase to ensure its clarity, coherence, and effectiveness. This step was crucial in identifying potential ambiguities or difficulties that respondents might face. To achieve this, we assembled a panel of eight participants, composed of both teachers and student teachers from different teaching networks, and asked them to complete the survey while providing detailed feedback on their experience.

Their recommendations led to several important modifications. Some questions were reworded to improve clarity, ensuring that respondents would interpret them correctly. We also adjusted the response formats of certain questions, replacing open-ended answers with multiple-choice options or Likert scales, when appropriate, to facilitate data analysis. This pre-test phase helped make the survey clearer, more reliable, and easier to use.

One of the key contributions came from Audrey Renson, who provided valuable recommendations for the clarity and organization of the survey. She suggested improving the introduction by making the text more structured, with clearer paragraphs and a larger font size to enhance readability. While we implemented a better layout for the introduction, we were unable to modify the font size due to the limitations of Google Forms. She also recommended changing the wording of a response option from “I prefer not to specify” to “I don’t want to answer this question”, as the latter phrasing is more neutral and respondent-friendly. Another crucial point she raised was ensuring that every question directly corresponded to a research question, which guided us to carefully review and refine the survey to ensure its alignment with our study’s objectives.

Lisa Svennberg, who works as an assistant in the Department of Didactics of Modern Languages, also provided constructive feedback, particularly regarding the multiple-choice options in the section about ChatGPT usage. She noted that she regularly used ChatGPT, but her specific way of using it was not represented in the predefined answer choices. This highlighted a limitation in the initial design of the questionnaire, leading me to expand the response options to better reflect the varied uses of ChatGPT in educational settings.

Similarly, Florence Van Hoof and Julie Vanhoof, both assistants in the same Department, suggested adding an “I don’t know” option to certain questions, as she found that, in some cases, she was unable to provide an informed response based on the available choices. However, it was not present in the options to avoid neutral responses. In addition, they recommended including an option stating “I didn’t learn to use it” for questions related to training and familiarity with ChatGPT, as some participants might not have received any formal instruction on AI tools. I took her suggestions into account and made these adjustments to improve the accuracy and inclusivity of the responses.

It is important to note that all of these women serve as mentor teachers, meaning they are not only teachers but also possess substantial expertise in language instruction.

These pre-test insights were invaluable in refining the survey. By incorporating these adjustments, we improved question clarity, response accuracy, and the user experience, ultimately enhancing the quality and reliability of the data we collected.

4.2.3 Distribution of the Survey

Once the survey had been finalized, we implemented a multi-channel distribution strategy to maximize participation and ensure that the responses reflected a diverse group of (future) language teachers. The questionnaire was officially launched on January 18, 2025 and remained open until early March. To maintain engagement and increase visibility, a reminder was shared on our social media platforms in mid-February. The survey was then shared through the following networks:

- Secondary school teachers: The survey was sent to in-service teachers in various schools, allowing us to gather insights from educators with different levels of experience and institutional backgrounds. Thanks to the support of Professor Simons and his team, we were able to reach a maximum number of teachers and schools.
- University students in teacher training programs: Given that part of our target population consisted of future teachers, we made sure to distribute the questionnaire among students enrolled in language education programs.
- Directors of higher education institutions: To further extend our reach, we contacted school directors, including the Haute École Les Rivageois and Haute École de la Ville de Liège, who helped circulate the survey among their students and teachers.
- Social media platforms: Recognizing the potential of online communities, we also posted the survey on Facebook, making it accessible to teachers who might not have been reached through institutional channels.

By using different ways to share the survey, we made sure it reached a wide range of people, including both teachers and future teachers. This variety gave us more useful and representative results, helping us better understand how ChatGPT is seen and used in language teaching.

4.2.4 Choosing the questions

In this section, I will explain why I chose to ask each question. Since the survey was in French and this dissertation is in English, I will translate every question to keep it uniformed. To

achieve this, I will justify each survey question by referring to the research (sub-)question it was designed to address.

The first part was about the respondents' profile.

Number	Questions from the survey	Research question(s) associated	Response methods	Goal
1	"You are: - A woman - A man - Non-binary - I do not want to answer this question"	RQ5-D: *Are there significant differences between men teachers and women teachers in terms of their attitudes towards ChatGPT?*	MCQ	Difference in usage depending on the genre.
2	"How old are you? - Less than 25 years old - Between 20 and 25 - Between 26 and 30 … until - More than 65"	RQ5-A: *Are there significant differences between older teachers and younger teachers in terms of their attitudes towards ChatGPT?*	MCQ	Difference in usage depending on the age.
3	"Are you: - A student - A teacher - A student and a teacher"	RQ5-B: *Are there significant differences between experienced teachers and students teachers in terms of their attitudes towards ChatGPT?*	MCQ	Difference in usage depending on the status.
4	If the respondent was a student: "What is your current level of study? - Bac1 - Bac2 - Bac3 - Master1 - Master2 - AESS - Master complementaire[25]"	RQ5-C: *Are there significant differences among students teachers from different levels of education in terms of their attitudes towards ChatGPT?*	MCQ	Difference in usage depending on their current level of study.
6	If the respondent was a teacher: "For how many years have you been teaching?" - Less than 1 year - Between 1 and 5 years - Between 6 and 10 years … until - More than 40 years"	RQ5-B: *Are there significant differences between experienced teachers and students teachers in terms of their attitudes towards ChatGPT?*	MCQ	Difference in usage depending on the teachers' experience.
7	"I'm teaching in: - DI - DS - Primaire - Haute école/Université - Promotion sociable - Autre"	RQ2-E: *Do (future) teachers' usage habits change depending on the type of audience?*	MCQ	Difference in usage depending on the type of audience.
9	"Which level of language do you teach? - L[26]1	RQ2-E: *Do (future) teachers' usage habits*	MCQ	Difference in usage depending

[25] This is a master for students who already graduated at the University and want to have a second degree.
[26] L1 = first foreign language, L2 = second foreign language; L3 = third foreign language.

	- L2 - L3"	*change depending on the type of audience?*		on the type of audience.
10	"I'm teaching in the network… - Libre - WBE - Communal officiel - Provincial officiel	RQ2-E: *Do (future) teachers' usage habits change depending on the type of audience?*	MCQ	Difference in usage depending on the type of audience.
11	"I'm teaching/I will teach: - English - Dutch - Spanish - Italian - German - French (as a foreign language)"	RQ2-D: *Do their usage habits change depending on the language taught?*	MCQ	Difference in usage depending on the language taught.

Table 7: Survey questions (part 1), their relation to research (sub-)questions and their objectives.

The second part was about Lisa Gayet's thesis. My part comes in the third place, which is about ChatGPT.

Number	Questions from the survey	Research question(s) associated	Response method	Goal
34	"Do you use ChatGPT in the private domain? - Yes - No"	RQ2-A: *To what extent do (future) langauge teachers in Liège currently use ChatGPT?*	MCQ	To observe if some teachers used it privately and not professionally, or vice versa.
35	"Do you use ChatGPT in the professional domain? - Yes - No"	RQ2-A: *To what extent do (future) language teachers in Liège currently use ChatGPT?*	MCQ	To observe if a lot of teachers use it already or if it is still new.
36	For those who answered 'No' to the Q11: "Why don't you use ChatGPT?" o "I'm not comfortable with Artificial Intellignce" o "I do not trust Artificial Intelligence" o "I am against Artificial Intelligence" o "I don't see how it could by useful in my teaching" o "I don't know well ChatGPT" o "I've never thought about it" o If the respondent was a student: "I think it is forbidden by the trainers/evaluators in the initial training" o If the respondent was a teacher: "I think it is forbiddden by the inspection" o "I think that everything should come from me and	RQ2-B: *What are the main reasons why (future) language teachers decide to use (or not to use) ChatGPT in their teaching practices?*	Likert scale format: "Strongly Disagree" – "Disagree" – "Agree" – "Strongly Agree"	To know the reasons why (future) language teachers decide not to use ChatGPT.

	to use ChatGPT is a form of plagiarism"			
37	For those who answered 'Yes' to the Q11: "Why do you use ChatGPT? o "ChatGPT gives me inspiration" o "ChatGPT corrects efficiently my lessons" o "ChatGPT participates efficiently in the correction of my students' works" o "ChatGPT creates good exercises" o "ChatGPT explains well the grammar" o "ChatGPT offers me qualitative documents for my reading comprehensions" o "ChatGPT offers me qualitative documents for my listening comprehensions"	RQ2-B: *What are the main reasons why (future) language teachers decide to use (or not to use) ChatGPT in their teaching practices* RQ6: *What specific tasks do (future) language teachers use ChatGPT for when preparing language lessons?*	Likert scale format: "Strongly Disagree" – "Disagree" – "Agree" – "Strongly Agree" – "I've never used it FOR that"	To know the reasons why (future) language teachers decide to use ChatGPT.
38	"If you use it in another way than those mentioned in the last question, please, precise it here:	RQ2-B: *What are the main reasons why (future) language teachers decide to use (or not to use) ChatGPT in their teaching practices?*	Open question (they could pass it if they didn't have anything to say).	To know the reasons why (future) language teachers decide to use ChatGPT.
39	"When do you use ChatGPT? o "To design or develop the lexical axis (vocabulary) of my courses" o "To design or develop the generic axis (textual genres) of my courses" o "To design or develop the grammatical axis (grammar) of my courses" o "To design or develop the functional axis (language functions) of my courses"	RQ1-C: *What specific language skills do teachers believe ChatGPT can support most effectively?*	Likert scale format: "Strongly Disagree" – "Disagree" – "Agree" – "Strongly Agree"	To observe if the respondents used it more for a specific axis of their classes.
40	"How often do you use ChatGPT? - For every lesson - For every didactic sequence - More or less once in the period - More or less once in the semester - More or less once a year."	RQ2-A: *To what extent do (future) language teachers in Liège currently use ChatGPT in their teaching practices?*	MCQ	To observe if the respondents used it regularly or not.
41	"When you use ChatGPT, generally you: - Take all the information offered	RQ3-C: *Do (future) language teachers use their critical thinking*	MCQ	To see if (future) teachers used it blindly by taking everything ChatGPT says or not.

	- Take a part of the information offered"	*skills when using ChatGPT?*		
42	"When you use ChatGPT, generally you: - Take all the information offered as it is presented - Modify the information offered"	RQ3-C: *Do (future) language teachers use their critical thinking skills when using ChatGPT?*	MCQ	To see if (future) teachers use their critical thinking skills when using it.
43	"If you modify the information suggested, why do you do so? o "To adapt the pedagogical content to the specific learning objectives of my didactic sequence" o "To adapt the content to the students' level" o "To correct mistakes of content (information)" o "To make the form of ChatGPT's responses more accessible to the students (ex: to simplify the vocabulary)" o "To add information (ex: to include examples)" o "To adapt the information based on the didactic frameworks	RQ3-C: *Do (future) language teachers use their critical thinking skills when using ChatGPT?*	Likert scale format: "Strongly Disagree" – "Disagree" – "Agree" – "Strongly Agree"	To understand the reason why they choose to modify the output produced by ChatGPT.
44	"Globally, do you think using ChatGPT contributes to improve the quality of your didactic sequences?"	RQ1-D: *Do (future) language teachers perceive ChatGPT as a tool than enhances creativity in lesson planning, or as something that limits their freedom to make pedagogical choices?*	Likert scale format: "Strongly Disagree" – "Disagree" – "Agree" – "Strongly Agree"	To observe if the respondents saw ChatGPT as beneficial or not.
45	"According to you, what is **the** added value of ChatGPT compared to units from manuals, sequences created by yourself and sequences encountered online?"	RQ1-D: *Do (future) language teachers perceive ChatGPT as a tool than enhances creativity in lesson planning, or as something that limits their freedom to make pedagogical choices?*	Open question.	To see if the opinions were diversified or if the majority thought the same. This was a good question to see clearly why and how ChatGPT is considered better than the other materials.
46	"According to you, what are the risks of using ChatGPT in the educational context?" o "Lack of information's reliability (content)" o "Negative impact on the teacher's creativity" o "Technological dependency: the teacher is no longer able to create a didactic sequence without ChatGPT"	RQ1-E: *What are the main fears or hesitations teachers have about introducing ChatGPT into the preparation of their language classes*	Likert scale format: "Strongly Disagree" – "Disagree" – "Agree" – "Strongly Agree"	To observe how strongly the respondents identified themselves with the sentences (to identify their main fears).

	○ "The first responses given (without any orientation from the teacher) could contain too many stereotypes" ○ "The information given by ChatGPT could not correspond to the expectations of the curricula" ○ "Risk of plagiarism" ○ "If you are a student: the approach does not correspond to the didactic models taught in the initial training (ex: PFE)" ○ "If you are a teacher: the approach does not correspond to the didactic models taught in the program"			
47	"Did you receive a training to use ChatGPT? (General use) - Yes - No"	RQ4-A: *Do (future) language teachers receive training on ChatGPT?*	MCQ	To see if trainings are frequent and to contrast it with the following question.
48	"Did you receive training to use ChatGPT in foreign languages? - Yes - No"	RQ4-A: *Do (future) language teachers receive training on ChatGPT?*	MCQ	To see if there was a difference with the previous question and to observe if it is present in schools or not.
49	"If you did receive training, you learned to use ChatGPT: - In the initial training - In the continuous training - Other"	RQ4-A: *Do (future) language teachers receive training on ChatGPT?*	MCQ	To see if there was a difference with the previous question and to observe if it is present in schools or not.
51	"If you didn't receive training, you learned to use ChatGPT: - From tutorials - By trial and error - Other" If the respondents chose "other", they were led to another question to precise it.	RQ4-A: *Do (future) language teachers receive training on ChatGPT?*	MCQ	To see if there was a difference with the previous question and to observe if it is present in schools or not.
53	"Do you think it would be interesting to develop trainings for the usage of ChatGPT for teachers in language education? - Yes - No"	RQ4-B: *Do (future) language teachers want to receive training on ChatGPT?*	MCQ	This question would serve as a proof that (future) teachers want to know how to use it in the professional domain.
54	"Which version of ChatGPT are you using currently or have you ever used? - The free version - The premium version - Both versions - None of these versions"	RQ2-F: *Do (future) language teachers use the premium version of ChatGPT?*	MCQ	To see if some of them were paying to receive responses from an upgraded version of ChatGPT.

55	"If you are using/ have used the premium version: what are the pedagogical or technical advantages that you observed in this version?".	RQ3-B: *According to them, is there any difference in the quality of responses given to premium users and free users?*	Open question	It enables me to understand the reasons why a teacher would want to pay more or less 22€ per month.
56	"Eventual comment" => the respondents were able to leave a comment if they wanted to add something.			

Table 8: Survey questions (part 2), their relation to the research (sub-)questions and their objectives.

4.3 Results

The survey was completed by 102 participants, providing a representative overview of language teachers and teacher trainees in Liège. Once the survey was closed, the responses were automatically compiled into an Excel file via Microsoft Forms, which provided a structured overview of the data. The platform also generated graphs for each question, which helped visualise general trends. To explore potential correlations, I manually sorted and filtered the data based on specific variables such as age, teaching experience, gender, or level of study. For example, to examine how age might influence perceptions of ChatGPT, I cross-referenced the participants' ages with their answers to a question related to whether they view ChatGPT as a beneficial tool in language education. This process involved identifying patterns (such as which age groups were more likely to see ChatGPT positively or negatively) and comparing response distributions across categories. Although the analysis was done manually, this approach allowed for a deeper understanding of how different profiles of (future) teachers perceive and use ChatGPT in their professional context.

4.3.1 Respondents' data

Gender distribution shows a strong female majority, with 83% identifying as women and 17% as men. This reflects a broader trend in the field of language education, where the profession remains largely female-dominated. This demographic imbalance is important to consider when analysing gender-based attitudes towards ChatGPT, as female perspectives are more heavily represented in the dataset.

In terms of age, respondents ranged from 20 to 55 years old, with the largest age group being 20–25 years (24%).

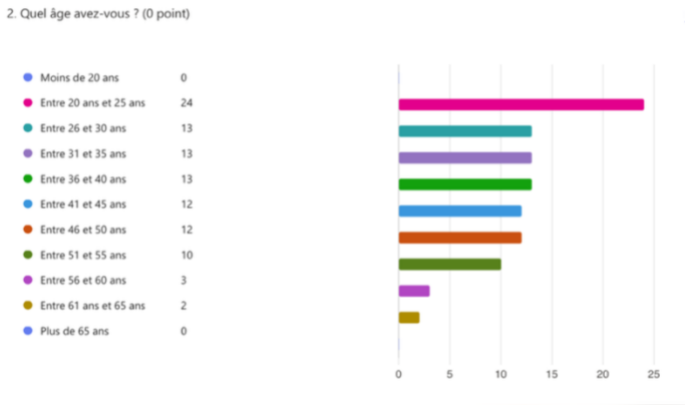


Figure 1: Survey – Question 2 (respondents' profile) – "How old are you?" – Results.

Regarding professional status, the majority of participants identified as in-service teachers (76%), while 20% were student teachers and 4% held both statuses. Among students, 35% were enrolled in a Master's degree, and 40% of student teachers were completing the AESS (Agrégation de l'Enseignement Secondaire Supérieur) program, indicating a strong academic background.

Participants also varied in teaching experience. The largest groups had between 1 and 5 years (17%), 16 and 20 years (17%), and 21 and 25 years (17%) of experience, followed closely by those with 11 to 15 years (16%), showing a balanced mix of early-career, mid-career, and late-career teachers.

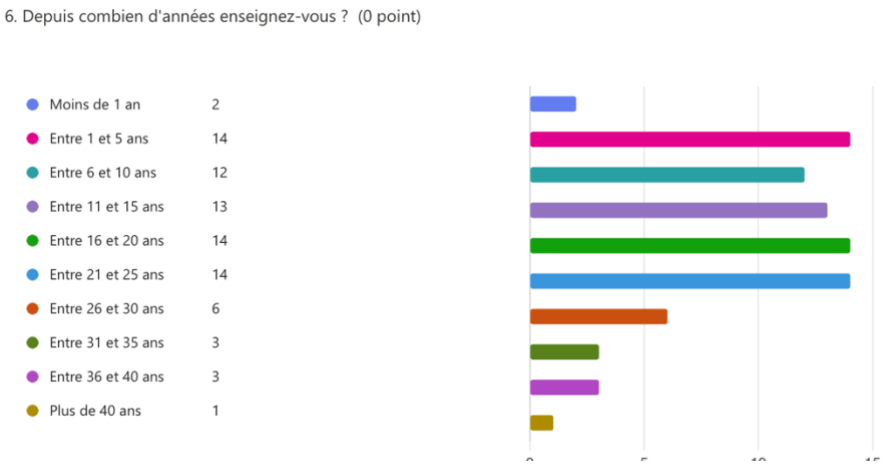


Figure 2: Survey – Question 6 (respondents' profile) – "How long have you been teaching?" – Results.

As for teaching levels, 66% of respondents taught in upper secondary education, 15% in lower secondary, and 7% in adult education ('promotion sociale'). This distribution highlights that most responses reflect the practices and needs of upper secondary language teaching.

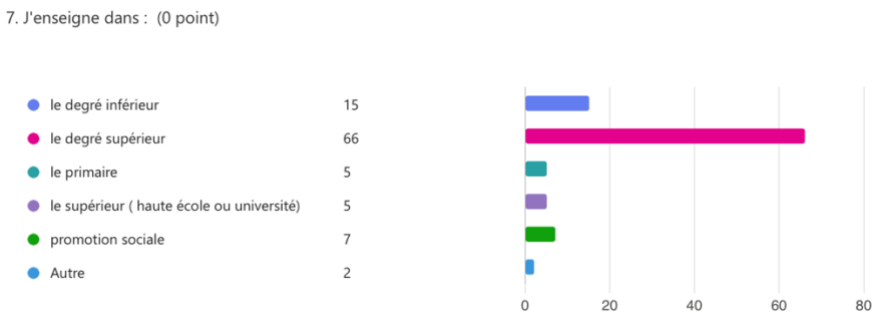


Figure 3: Survey – Question 7 (respondents' profile) – "I teach in:" – Results.

Regarding language profiles, 51% of respondents taught a first foreign language (L1), 38% a second foreign language (L2), and 12% a third (L3), showing the multilingual nature of language instruction in Belgium.

In terms of school networks, 55% of the participants were employed in the free subsidised network, followed by 27% in the official system (WBE), indicating that responses are primarily drawn from these two networks.

Finally, participants represented a range of language specialisations, with English being the most commonly taught language (49%), followed by Dutch (25%), Spanish (12%), German (11%), French (2%), and Italian (1%). This distribution highlights the dominant position of English in foreign language instruction in the region.

4.3.2 The main survey

The data analysed in this section corresponds exclusively to the third part of the questionnaire, which was designed to investigate the usage patterns, attitudes, and opinions regarding the integration of ChatGPT into language teaching practices. The findings are structured around the research questions (each section corresponds to a research question and its sub-questions associated) and include descriptive insights as well as preliminary comparative trends across age, experience, and gender.

1) Responses to RQ2 : *What do (future) language teachers in Liège say about their use of ChatGPT?*

RQ2-A: *To what extent do (future) language teachers in Liège currently use ChatGPT in their teaching practices?*

Out of the 102 participants who completed the questionnaire, a significant majority indicated that they currently use ChatGPT in their professional practice as language teachers or future teachers. Specifically, 70 respondents (69%) affirmed that they use ChatGPT in a professional context, while 32 respondents (31%) reported not using it in their teaching activities. This finding shows that ChatGPT is already quite popular among (future) language teachers in Liège, even though it is still a fairly new tool in education. However, this high percentage should be interpreted with caution, since the people who answered the survey chose to take part and were likely already interested or familiar with ChatGPT.

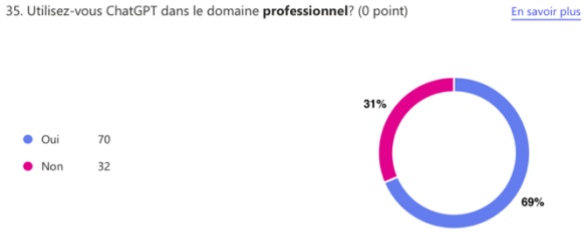


Figure 4: Survey – Question 35 – "Do you use ChatGPT in the professional domain?" – Results.

In addition, out of the 70 participants who reported using ChatGPT in their professional teaching practices, 58 also reported using it in their private lives. Only 12 respondents reported using it professionally without using it privately. In contrast, among the 32 participants who do not use ChatGPT professionally, only 3 reported using it privately. The vast majority of them, 29 participants, neither use ChatGPT professionally nor privately.

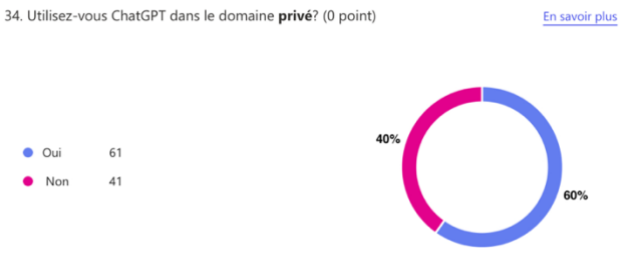


Figure 5: Survey – Question 34 – "Do you use ChatGPT in the private domain?" – Results.

Conclusion to RQ2-A: These numbers show that people who use ChatGPT in their personal life are more likely to use it at work too. Being comfortable with the tool privately often leads to using it in teaching. On the other hand, those who do not use it for work, usually do not use it at all. Other factors, like age, will be looked at later to better understand what influences how ChatGPT is used.

RQ2-B: *What are the main reasons why (future) language teachers decide to use (or not to use) ChatGPT in their teaching practices?*

The reasons underlying the decision to either adopt or reject ChatGPT in teaching practices were varied, revealing a nuanced perception of the tool. Among 70 respondents who use ChatGPT professionally, the most common reason cited was its capacity to serve as a source of inspiration. When asked whether ChatGPT gave them inspiration for lesson design, 88.6% responded positively: 37 (52.9%) selected "Strongly agree", and 25 (35.7%) "Agree". Only 4 (5.7%) disagreed with this statement, and a mere 1 person strongly disagreed. An additional 3 respondents specified that they had never used ChatGPT for that specific purpose.

Figure 6: Survey – Question 37 – "Why do you use ChatGPT?" – Results

On the contrary, among the 32 respondents who reported not using ChatGPT professionally, the reasons for non-use included a lack of familiarity and ethical concerns. When asked whether they did not know ChatGPT well enough, 12 (37.5%) selected "Strongly agree", and 13 (40.6%) chose "Agree". 5 respondents disagreed with the statement, and only 2 participants strongly

disagreed, suggesting that unfamiliarity with the tool is a common but not universal reason for abstention.

Another key reason for not using ChatGPT is a perceived threat to the teacher's pedagogical integrity or authorship. In response to the statement, "I believe everything should come from me, and using ChatGPT is a form of plagiarism", 6 respondents (19,4%) strongly agreed, and 7 (22,6%) agreed. Meanwhile, 15 (48,4%) disagreed, and 3 (9,7%) strongly disagreed.

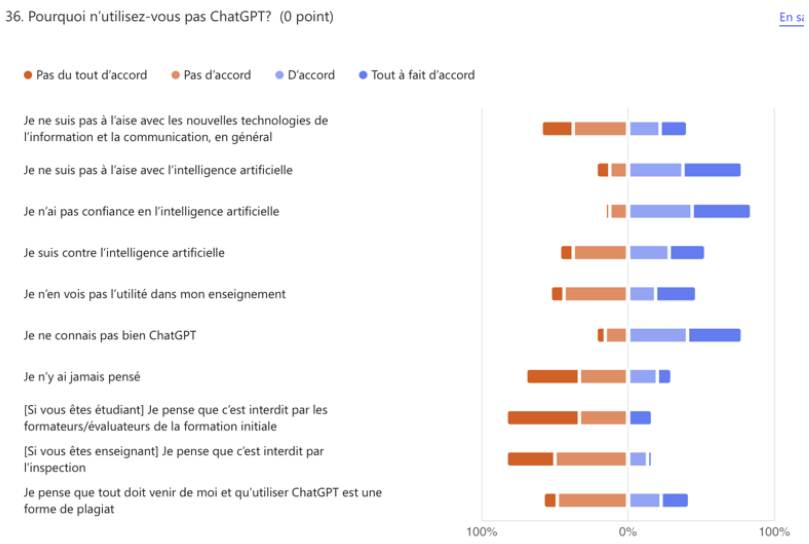


Figure 7: Survey – Question 36 – "Why do you not use ChatGPT?" – Results.

Conclusion to RQ2-B: These findings underscore the diversity of professional attitudes towards ChatGPT. Many see it as a helpful tool for creativity, while others feel it does not match their teaching values or are unsure how to use it due to lack of confidence or understanding. This mix of excitement and hesitation shows the need for training and open discussion on how to use ChatGPT in a smart and ethical way in language teaching.

RQ2-C: *What specific tasks do (future) language teachers use ChatGPT for when preparing language lessons?*

The results of the survey highlight the broad range of pedagogical tasks for which (future) language teachers in Liège employ ChatGPT (see figure 7). These tasks span both sequence design and instructional support.

A) The creation of exercises

One commonly cited use of ChatGPT regards the creation of exercises. When asked whether "ChatGPT creates good exercises", a total of 55 out of 70 respondents (78.6%) expressed a positive opinion: 37 selected "Agree" (52.9%) and 18 selected "Strongly agree" (25.7%). Meanwhile, 9 respondents (12.8%) gave a negative opinion, with 8 selecting "Disagree" (11.4%) and 1 selecting "Strongly disagree" (1.4%). Finally, 6 respondents (8.6%) said they had never used ChatGPT for that specific purpose.

B) Grammar explanation

Another relevant task is grammar explanation. In this area, 33 respondents out of 70 (47,2%) expressed a positive opinion, with 30 (42.9%) agreeing that ChatGPT explains grammar well, while 3 (4.3%) strongly agreed. However, 12 respondents (17,1%) expressed a negative opinion: 8 (11.4%) disagreed, 4 (5.7%) strongly disagreed, and 25 (35.7%) stated that they had never used ChatGPT to explain grammar.

C) Lesson correction

ChatGPT is also used by some teachers for lesson correction. When asked whether ChatGPT effectively corrects their own lessons, 11 teachers out of 70 (15.7%) showed a positive view: 10 (14.3%) agreed and 1 (1.4%) strongly agreed. However, 15 (21.4%) disagreed, and 9 (12.9%) strongly disagreed, which shows that 24 respondents (34.3%) expressed a negative opinion. A notable 35 respondents (50%) claimed that they had never used ChatGPT for this purpose, suggesting that this specific application is either not well known or not considered reliable by a large part of the teachers.

D) Assignments' correction

Similarly, teachers were asked if ChatGPT assists in correcting student work. In this case, 9 (12.9%) agreed and 12 (17.1%) strongly disagreed, while 7 (10%) disagreed. A large majority - A total of 42 participants (60%) - said they had never used ChatGPT for this purpose, making it the least common task among those listed. This result may be due to concerns about accuracy, evaluation ethics, or the complexity of assessing students' work via AI.

Conclusion to RQ2-C: These findings confirm that while the most popular use cases for ChatGPT are related to content generation (such as exercises or grammar explanations), tasks

involving evaluation or correction are approached with more caution or remain underexplored by the teachers.

RQ2-D: *Do (future) teachers' usage habits change depending on the language taught?*

Although the survey did not explicitly ask respondents to identify which language(s) they taught in each ChatGPT-related question, some conclusions can be drawn from other data. Participants had the option to select multiple languages (e.g., English and Dutch, or English and Spanish), and their answers were grouped accordingly to see if ChatGPT use differed. This allowed to investigate whether usage habits regarding ChatGPT varied according to the language taught.

The analysis revealed that how often ChatGPT is used at work depends on the language taught. Out of 85 responses related to teaching English, 62 (72.9%) reported using ChatGPT in their professional practice, while 23 (27,1%) said they did not. This makes English the most represented language in terms of ChatGPT adoption. Similarly, among German teachers (20 participants), 70% reported professional use, while 30% did not.

Spanish teachers (21 participants) followed closely, with 66.7% indicating professional use. Dutch teachers (43 respondents), on the other hand, demonstrated a slightly lower level of engagement, with 62.8% using ChatGPT professionally and 37.2% not using it in that context. While these differences are not extremely large, they suggest that teachers' willingness to adopt AI tools may indeed vary depending on the linguistic context.

It is worth noting that teachers of French as a foreign language all reported using ChatGPT professionally (100%), although the group was small (3 respondents). Similarly, among the very limited number of Italian teachers (2 participants), half reported using ChatGPT professionally.

Conclusion to RQ2-D: These findings suggest that the language taught can affect how likely teachers are to use ChatGPT. Since ChatGPT was mainly trained in English, English teachers may find the tool more useful and accurate, which could explain why they use it more. In other languages, the output may be less helpful, so teachers use it more carefully. These differences

may also come from teaching styles or curriculum expectations in different language departments.

RQ2-E: *Do (future) teachers' usage habits change depending on the type of audience?*

The survey reveals a nuanced relationship between teachers' usage habits and the audience they teach, including factors such as school level, language level, and educational network.

A) Educational level:

Educational level	Each lesson	Each didactic sequence	Once per period	Once per semester	Once per year
"Degré supérieur"	6	18	23	5	2
"Degré supérieur" + "degré inférieur"	0	4	2	1	0
"Degré supérieur" + "Université/Haute école"	0	1	0	0	1
"Degré inférieur"	0	1	3	3	0
"Promotion sociale"	0	3	2	0	0
"Primaire"	1	0	2	1	0
"Université/Haute école"	0	1	0	0	0
"Université/Haute école" + "promotion sociale"	0	0	1	0	0

Table 9: Cross-analysis of educational level and frequency of ChatGPT use among respondents.

To answer this question, responses were cross analysed based on two variables: the educational level taught and the frequency of ChatGPT use. The following usage categories were considered: use of ChatGPT for each lesson, each didactic sequence, once per period, once per semester, and once per year.

Among the 54 respondents teaching in the "degré supérieur", usage of ChatGPT is clearly more frequent: 6 use it for each lesson, and 18 for each didactic sequence. This makes "degré supérieur" the group with the highest reported regular use. A similar trend was observed among the 7 respondents teaching in both the "degré supérieur" and "degré inférieur": 4 of them declared using ChatGPT for each didactic sequence.

In contrast, among the 7 respondents who teach only in the "degré inférieur", none reported using ChatGPT for each lesson, and only one mentioned using it for a sequence. Primary school teachers (4 respondents) also showed minimal use: only one reported using it for each lesson,

and none per sequence. Among the 5 respondents working in “promotion sociale”, 3 reported using ChatGPT per sequence, while none used it for every lesson.

For the teachers in higher education (Haute Ecole/Université), the distribution is uneven and the sample size is too small to support any reliable information. For this reason, their responses have not been analysed in detail, as the limited sample size would not allow for meaningful conclusions.

B) School network:

School network	Each lesson	Each didactic sequence	Once per period	Once per semester	Once per year
WBE (22 out of 70)	0	5	11	4	2
Libre (36 out of 70)	6	12	15	3	0
Communal (11 out of 70)	0	5	4	1	1
Provincial (4 out of 70)	0	2	1	1	0

Table 10: Cross-analysis of school network and frequency of ChatGPT use among respondents.

The analysis reveals that ChatGPT usage habits vary across different school networks. In the “Libre” network (free subsidised network), usage is both higher and more frequent: 16.7% of teachers (6 out of 36) use ChatGPT for each lesson, and 33.3% use it for each didactic sequence (12 out of 36). In contrast, no teachers in the WBE, Communal, or Provincial networks reported using it for every lesson.

In the Communal network, 45.5% of respondents use ChatGPT for each didactic sequence (5 out of 11), the highest proportion across all networks for that category. Similarly, 50% of teachers in the Provincial network use it per sequence, although the sample for that group is very small (4 participants).

In the WBE network, the most common use is once per period (50%), with only 22.7% using it per sequence (5 out of 22) and none for each lesson. This suggests a more moderate or occasional use compared to the Libre or Communal networks.

Conclusion to RQ2-E: Although the data suggests that teachers working in the “degré supérieur” and those teaching in the “Libre” and Communal networks tend to use ChatGPT more regularly in their lesson planning, these trends should be interpreted with caution. The survey responses were not evenly distributed: over half of the participants (54 out of 102) teach in the “degré supérieur”, while the Communal and Provincial networks had fewer respondents.

This imbalance may have influenced the results, making ChatGPT usage appear more frequent in certain contexts. To draw more reliable conclusions, a more balanced sample across different educational levels and school networks would be needed. It would also be worth exploring whether factors such as access to training, institutional support, or specific pedagogical expectations contribute to these differences in how ChatGPT is used in classrooms.

RQ2-F will be answered later with the RQ3-B.

2) **RQ1:** ***What are the (perceived) pedagogical benefits and limitations of using ChatGPT in language education?***

RQ1-C: *What specific language skills do teachers believe ChatGPT can support most effectively?*

Respondents were also asked about their perception of ChatGPT's effectiveness in supporting specific language skills, particularly reading and listening comprehension.

Regarding reading comprehension, on the one hand, 41 out of 70 respondents (58.5%) expressed a positive view: 36 (51.4%) agreed that ChatGPT provides quality materials, while 5 (7.1%) strongly agreed. On the other hand, 13 respondents (18.5%) showed a negative opinion: 8 participants (11.4%) disagreed, 5 (7.1%) strongly disagreed. The remaining 16 participants (22.9%) stated that they had never used ChatGPT for this task.

In the case of listening comprehension, the perception of effectiveness was slightly lower. Only 11 respondents (15.7%) expressed a positive view, with 10 (14.3%) who selected agreed, and 1 (1.4%) who strongly agreed that ChatGPT can provide quality audio comprehension materials. In comparison, 21 participants (30%) responded negatively: 14 (20%) disagreed, 7 (10%) strongly disagreed. A majority of respondents, 38 people (54.3%), declared never having used ChatGPT for this task.

Conclusion to RQ1-C: The relatively lower use of ChatGPT for listening comprehension may reflect technical limitations, such as the tool's text-based nature or the need for external resources to create authentic audio input. Nevertheless, the tool is perceived as a valuable asset

for reading tasks, possibly due to its capacity to generate varied and customizable texts that align with students' proficiency levels and curricular goals.

These findings show that people mostly use and trust ChatGPT to help with receptive skills, especially reading, as well as with creating learning materials and teaching grammar. However, it is used less for tasks that involve speaking, listening, or evaluation, either because people are not sure how well it works for these things, or they do not know it can be used that wa

RQ1-D: *Do (future) language teachers perceive ChatGPT as a tool that enhances creativity in lesson planning, or as something that limits their freedom to make pedagogical choices?*

The perception of ChatGPT's impact on the quality and creativity of didactic sequences varied among respondents, yet the majority viewed it as a beneficial tool. When asked whether they agreed with the statement "ChatGPT contributes to improving the quality of my lesson sequences", 60 out of 102 respondents (58.8%) expressed a positive opinion: 46 (45.1%) answered "Agree", and 14 (13.7%) responded with "Strongly agree". In contrast, 42 participants (41.2%) showed a more critical view: 27 respondents (26.5%) selected "Disagree", while 15 (14.7%) answered "Strongly disagree".

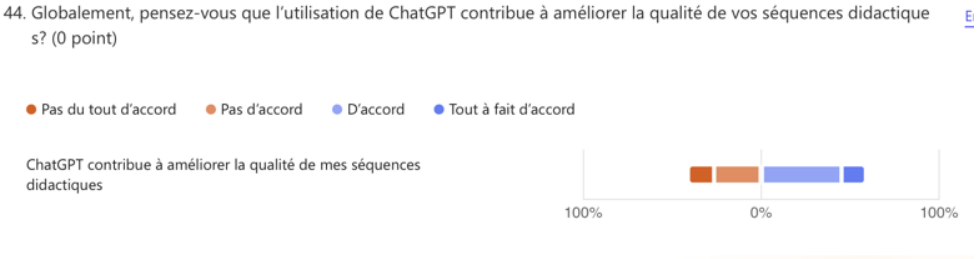


Figure 8: Survey – Question 44 – "Globally, do you think ChatGPT contributes to improving the quality of your didactic sequences?" – Results.

Conclusion to RQ1-D: These numbers indicate that 58.8% of the participants have a positive view of ChatGPT in this regard, while 41.2% hold more critical or negative perspectives. Their concerns may come from ethical issues, doubts about content quality, or discomfort with using AI in teaching. The results suggest the real divide is not just between users and non-users, but between those who see ChatGPT as a creative tool and those who are more careful, worrying

it might reduce their control as teachers or lead to overuse. Other variables, like age, will be analysed later to better understand what influences how ChatGPT is used.

RQ1-E: *What are the main fears or hesitations teachers have about introducing ChatGPT into the preparation of their language classes?*

The questionnaire also explored perceived risks associated with the integration of ChatGPT. Three key concerns emerged from the data.

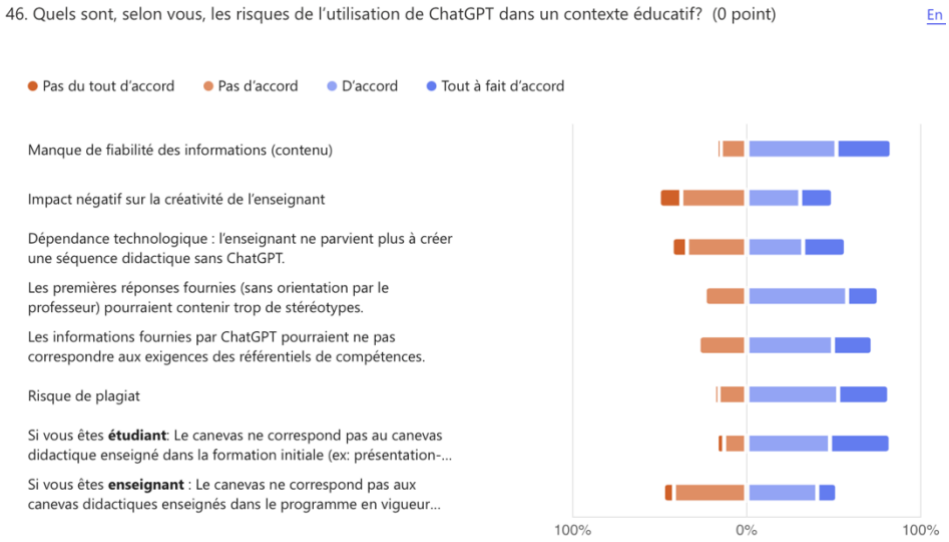


Figure 9: Survey – Question 46 – "According to you, what are the risks of using ChatGPT in education?" – Results.

The first concern was the risk of technological dependence. When asked whether they agreed with the statement "Teachers may no longer be able to create didactic sequences without ChatGPT", 58 out of 102 participants (56.9%) expressed some level of agreement: 25 (24.5%) selected "Strongly agree", and 33 (32.4%) selected "Agree". In contrast, 44 respondents (43.1%) showed a more skeptical view, with 35 (34.3%) who disagreed, and 9 (8.8%) who strongly disagreed. These results suggest that over half of the participants recognise a potential risk of becoming too reliant on ChatGPT for lesson planning.

The second concern focused on alignment with curricular standards. 73 respondents (71.5%) expressed concern: 50 (49.0%) agreed that the information provided by ChatGPT might not correspond to official curriculum requirements, and 23 (22.5%) strongly agreed. Only 28 respondents (27.5%) disagreed with this concern.

The third concern was the risk of plagiarism. In response to the statement "Using ChatGPT in lesson planning is a form of plagiarism", opinions were more divided, though negative views dominated. 32 respondents (31.5%) expressed concern: 23 (22.6%) selected "Agree" and 9 (19.4%) "Strongly agree". In contrast, 59 participants (58.1%) rejected this view: 49 (48.4%) chose "Disagree" and 10 (9.7%) "Strongly disagree". This indicates that while the majority do not associate ChatGPT use with plagiarism, a significant minority still carry ethical concerns about authorship and originality.

Conclusion to RQ1-E: These numbers suggest that even among users of ChatGPT, many approach its output with caution and a sense of professional responsibility to validate or adapt AI-generated materials before use. This observation reinforces the relevance and necessity of the present study.

3) RQ5: *How do personal and contextual factors influence attitudes and usage of ChatGPT?*

RQ5-A: *Are there significant differences between older teachers and younger teachers in terms of their attitudes towards ChatGPT?*

To explore this question, participants were asked: "Globally, do you think ChatGPT contributes to improving the quality of your didactic sequences?". The responses, when analysed by age group, reveal a clear generational divide in attitudes.

Age group	Strongly agree	Agree	Disagree	Strongly disagree
Between 20-25	7	13	2	2
Between 26-30	3	6	3	0
Between 31-35	0	8	4	1
Between 36-40	1	7	2	2
Between 41-45	1	4	6	1
Between 46-50	1	3	6	2
Between 51-55	1	4	2	2
Between 56-60	0	0	0	3
Between 61-65	0	0	0	1

Table 11: Distribution of the age of the respondents and their degree of agreement to the question "Globally, do you think Chatgpt contributes to improving the quality of your didactic sequences?".

Among teachers aged 20–25 (24 respondents), a majority expressed a positive opinion: 13 selected "Agree" and 7 "Strongly agree", meaning that 83% of this group viewed ChatGPT as a beneficial tool. Only 2 respondents disagreed, and none selected "Strongly disagree".

Similarly, in the 26–30 age group (13 respondents), 10 participants expressed positive views (7 "Agree" and 3 "Strongly agree"), with very few negative responses.

In contrast, older participants showed more hesitation or rejection of the tool. Among those aged 46–50 (12 respondents), 8 expressed a negative opinion (6 "Disagree" and 2 "Strongly disagree"), while only 3 agreed with the statement. This pattern continues in the 51–55 age group (10 respondents), where most responses were also negative. In the oldest categories—56–60 (3 respondents) and 61–65 (2 respondents)—participants overwhelmingly chose "Disagree" or "Strongly disagree".

Conclusion to RQ5-A: These results suggest a strong generational trend: younger teachers tend to view ChatGPT as a useful support for lesson planning, while older teachers are generally more critical or resistant. This difference may be linked to varying levels of digital literacy, openness to new technologies, or a stronger attachment to traditional teaching methods among older respondents.

RQ5-B: *Are there significant differences between experienced teachers and student teachers in terms of their attitudes towards ChatGPT?*

The survey results show clear differences in how ChatGPT is perceived depending on teaching experience. To answer this question, results were analysed in the same way as for the previous research question.

Years of teaching experience	Strongly agree	Agree	Disagree	Strongly disagree
Less than a year	1	1	0	0
Between 1-5	4	4	2	4
Between 6-10	4	3	2	3
Between 11-15	2	5	4	2
Between 15-20	2	5	4	3
Between 21-25	1	4	8	1
Between 26-30	0	4	2	0
Between 31-35	0	2	1	0
Between 36-40	0	3	0	0
More than 40	0	1	0	0

Table 12: Distribution of teachers' experience and their degree of agreement to the question "Globally, do you think Chatgpt contributes to improving the quality of your didactic sequences?".

Among the least experienced respondents, those with less than one year of experience (2 respondents), one strongly agreed, and one agreed with the statement. In the group with 1 to 5 years of experience (14 respondents), 4 strongly agreed, 4 agreed, 2 disagreed, and 4 strongly

disagreed. Although the group showed a wide distribution, 57% (8 out of 14) expressed agreement to some extent, while 43% (6 out of 14) expressed disagreement.

Among those with 6 to 10 years of experience (12 respondents), 4 strongly agreed, 3 agreed, 2 disagreed, and 3 strongly disagreed. Again, 58.3% (7 out of 12) had a positive view, while 41.7% (5 out of 12) disagreed to some extent. This suggests that teachers in the early and mid-career stages show an interest in ChatGPT.

In contrast, teachers with 11 to 15 years of experience (13 respondents) showed a more divided attitude: only 2 strongly agreed and 5 agreed, but 4 disagreed and 2 strongly disagreed. This means 53.8% (7 out of 13) agreed overall, while 46.2% (6 out of 13) disagreed. Teachers with 15 to 20 years of experience (14 respondents) showed a balanced attitude: 2 strongly agreed, 5 agreed, 4 disagreed, and 3 strongly disagreed. These results show that 50% expressed agreement (7 out of 14) and 50% disagreement (7 out of 14).

The most critical group appeared to be teachers with 21 to 25 years of experience (14 respondents): only 1 strongly agreed and 4 agreed, while 8 disagreed and 1 strongly disagreed. Here, 35.7% (5 out of 14) expressed agreement, while a significant 64.3% (9 out of 14) expressed disagreement. This group showed the highest rate of disagreement in the entire sample.

Among the more experienced teachers, attitudes towards ChatGPT were generally positive, though more measured. In the group with 26 to 30 years of experience (6 respondents), 4 agreed and 2 disagreed with the statement, resulting in an agreement rate of 66.7%. Teachers with 31 to 35 years of experience (3 respondents) also leaned towards agreement, with 2 agreeing and 1 disagreeing. Similarly, all 3 respondents in the 36 to 40 years bracket agreed with the statement, although none expressed strong agreement. The only respondent with over 40 years of experience also agreed.

While the number of participants in these more senior groups was relatively small, the results suggest a cautious openness to ChatGPT among more seasoned teachers, even if they are less likely to show strong enthusiasm.

Conclusion to RQ5-B: In summary, the data indicates a trend where less experienced teachers are more likely to strongly agree with the usefulness of ChatGPT, while more experienced teachers tend to agree moderately or express disagreement. The group with 21 to 25 years of experience was the most critical, and no teachers with more than 25 years of experience strongly agreed with the positive statement. These results suggest that years of experience may influence how open teachers are to integrating new technologies like ChatGPT into their practice.

RQ5-C: *Are there significant differences among student teachers from different levels of education (Bac 1-2-3, Master 1-2, AESS, Master complémentaire) in terms of their attitudes towards ChatGPT?*

To provide a more detailed response to this research question, the data was analysed by separating respondents into two categories based on their professional status: those who are currently student teachers and, those who are already working as in-service teachers. These two groups were identified based on the version of the question "What is your level of education?" they responded to in the survey. The analysis then focused on how these participants answered the statement "ChatGPT contributes to improving the quality of my didactic sequences". Their responses were then examined in relation to their level of education.

Among the student teachers (20 respondents), the majority of responses showed a positive perception of ChatGPT as a useful tool for lesson planning. Specifically, 12 respondents (60%) selected the option "Agree" and 5 respondents (25%) selected "Strongly agree". This means that 17 out of 20 student teachers (85%) expressed a positive opinion. On the other hand, two respondents (10%) selected "Disagree" and one respondent (5%) selected "Strongly disagree", indicating a small minority with a negative view.
When we look more closely at the different levels of education within the student teacher group, we observe the following details:

Level of education	Strongly disagree	Disagree	Agree	Strongly agree
Bac 1	0	0	0	0
Bac 2	0	1	0	1
Bac 3	0	0	2	1
Master 1	0	0	2	1
Master 2	0	0	5	2
Master complémentaire	0	0	2	0
AESS (students)	1	1	0	0
In-service student teachers	0	1	2	1

Table 13: Distribution of students' educational level and their degree of agreement to the question "Globally, do you think Chatgpt contributes to improving the quality of your didactic sequences?".

By contrast, the group of in-service teachers (4 respondents) showed a slightly more reserved attitude. Two respondents (50%) selected "Agree", and one respondent (25%) selected "Strongly agree". These results indicate that 3 out of 4 teachers expressed a positive view of ChatGPT. However, one teacher (25%) selected "Disagree", which suggests that there is some hesitation in this group compared to the student teachers. None of the in-service teachers selected "Strongly disagree", which implies that while some concerns exist, strong rejection of ChatGPT was not present among these four participants. Two of these four in-service teachers reported following the AESS programme, and both expressed negative views. The two remaining teachers held master's degrees and were more favourable towards the use of ChatGPT.

However, it is important to interpret the results for in-service teachers with caution. Given the very small sample size (4 participants), each response significantly affects the percentages. For example, one person represents 25% of the group. As a result, these findings are not statistically reliable and cannot be generalised to the broader population of language teachers in the French-speaking Community of Belgium. They should be seen as exploratory insights rather than firm conclusions.

Conclusion to RQ5-C: This analysis shows that student teachers, especially those enrolled in Master's programmes, tend to view ChatGPT more positively than others. They are more likely to perceive it as a useful tool for improving the quality of their didactic sequences. On the other hand, participants from AESS programmes, whether students or in-service teachers, appear more hesitant or critical. While a similar trend of openness can be observed among some in-service teachers with Master's-level enrolment, the small number of responses prevents any definitive interpretation. Further research with a larger sample would be necessary to confirm these patterns.

RQ5-D: *Are there significant differences between men and women teachers in terms of their attitudes towards ChatGPT?*

Gender	Strongly disagree	Disagree	Agree	Strongly agree
Men	2	3	7	5
Women	13	24	39	9

Table 14: Distribution of gender and their degree of agreement to the question "Globally, do you think ChatGPT contributes to improving the quality of your didactic sequences?".

To investigate whether gender plays a role in how language teachers perceive the value of ChatGPT, responses to the statement "ChatGPT contributes to improving the quality of my didactic sequences" were analysed according to the gender declared by participants. In total, 17 male and 85 female respondents answered this question.

Among male respondents, the majority expressed a positive opinion: 7 selected "Agree" and 5 selected "Strongly agree", meaning that 70.6% of men saw ChatGPT as a beneficial tool. The remaining 29.4% expressed disagreement, with 3 selecting "Disagree" and 2 selecting "Strongly disagree".

Among female respondents, 39 agreed and 9 selected strongly agreed, which amounts to 56.5% expressing a positive view. Meanwhile, 24 selected "Disagree" and 13 "Strongly disagree", resulting in 43.5% of women expressing a more negative attitude.

Although male respondents appeared more favourable overall toward the use of ChatGPT, the difference between the two groups (70.6% positive among men versus 56.5% among women) is relatively modest. Moreover, due to the small number of male participants (n = 17), even one or two answers could significantly influence the overall proportions for that group. As such, this difference should be interpreted with caution and cannot be considered a clear contrast. Nevertheless, according to the statistics from the Wallonia-Brussels Federation, the data confirms that women are more represented among teachers than men. The overall feminsiation rate in 2024 is 73.1% (WBF, n.d).

Conclusion to RQ5-D: While some differences appear between the responses of male and female participants, the sample size for men is too small to draw reliable conclusions. The gap between the two groups is moderate and could easily be influenced by the small number of

male respondents. Therefore, based on this dataset, there is no strong evidence of a correlation between gender and attitudes toward ChatGPT. The overall results suggest that perceptions of ChatGPT's usefulness are relatively balanced across genders in this sample.

4) **RQ4:** ***What training do (future) language teachers receive or expect regarding ChatGPT?***

RQ4-A: *Do (future) language teachers receive training on ChatGPT?*

The majority of participants reported having received no formal training in the use of ChatGPT, whether for general or pedagogical purposes. In response to the question "Have you received training in the general use of ChatGPT?", 88 respondents out of 102 (86.3%) answered "No", and only 14 (13.7%) answered "Yes". When asked if they had received training specifically for using ChatGPT in foreign language teaching, the results were even more decisive: 97 respondents (95.1%) answered "No", and only 5 respondents (4.9%) answered "Yes".

For respondents who answered that they had not received training, a follow-up question asked how they had learned to use ChatGPT. Sixty participants (58.8%) selected "through trial and error", while 22 (21.6%) selected "Other". Five participants reported combining trial-and-error with other forms of learning, and a small number (six in total) mentioned learning from tutorials, either alone or in combination with other approaches.

Conclusion to RQ4-A: These findings show that most teachers have learned to use ChatGPT on their own. This raises questions about how well they understand it and whether they are thinking about the ethical and teaching-related issues when using AI in class. It is clear that teacher training programs and professional development have an opportunity to step in and offer more support.

RQ4-B: *Do (future) language teachers want to receive training on ChatGPT?*

The majority of respondents expressed a clear interest in receiving formal training on the use of ChatGPT for language teaching. In response to the question "Do you think it would be useful

to develop training sessions on the use of ChatGPT for language teachers?", 92 participants out of 102 responded "Yes", while only 10 responded "No". This means that 90.2% of the sample showed a strong interest in pedagogical training on this tool.

Conclusion to RQ4-B: This strong demand is consistent with the findings discussed earlier in RQ4-A, where most teachers indicated they had not received any formal training and had instead taught themselves how to use ChatGPT. It reinforces the conclusion that even among current users, there is a desire for structured guidance, both to use the tool more effectively and to address concerns such as content quality, alignment with curricular goals, and ethical considerations.

5) **RQ3:** ***How do (future) language teachers perceive and evaluate the quality of ChatGPT-generated content?***

RQ3-A: *Which criteria do teachers apply to evaluate the quality of ChatGPT-generated content before using it in class?*

One of the final items in the survey asked respondents whether they change ChatGPT's output before using it in class.

The reasons most frequently cited for modifying the content include adapting the material to match the learning objectives of their sequence. In total, a majority of 63 teachers out of 67 (94.1%) expressed agreement to this statement: 31 strongly agreed (46.3%) and 32 agreed (47.8%). This motivation received the highest degree of agreement across the response scale. More or less the same was observed for adjusting the level of difficulty to suit their students: 58 out of 67 participants (86.6%) showed agreement, with 27 (40.3%) who strongly agreed and 31 (46.3%) who agreed.

Respondents also agreed for the need to insert concrete examples: a majority of 60 (future) teachers agreed with the statement, with 28 (41.8%) opting for "Strongly agree" and 32 (47.8%) for "Agree".

Moroever, 58 participants (86.6%) modify ChatGPT's output to restructure the information to improve readability. Indeed, 31 (46.3%) strongly agreed and 27 (40.3%) agreed with the statement.

Other mentioned reasons were the correction of factual or linguistic errors, making the language more accessible (e.g., simplifying vocabulary). A majority of respondents, 82.1% (57 out of 67), agreed with this statement. Specifically, 18 (26.9%) selected "Strongly Agree" and 37 (55.2%) selected "Agree". Also, some respondents change the output produced by ChatGPT to adapt it to the different didactic models (77.6%). 21 (31.3%) strongly agreed and 31 (46.3%) agreed with the statement.

Finally, the statement that received the highest percentage of disagreement was the one suggesting that teachers modify ChatGPT's output to correct errors in the information provided: 23 participants (34.3%) disagreed with this statement, with 7 who strongly disagreed (10.4%) and 16 (23.9%) who disagreed.

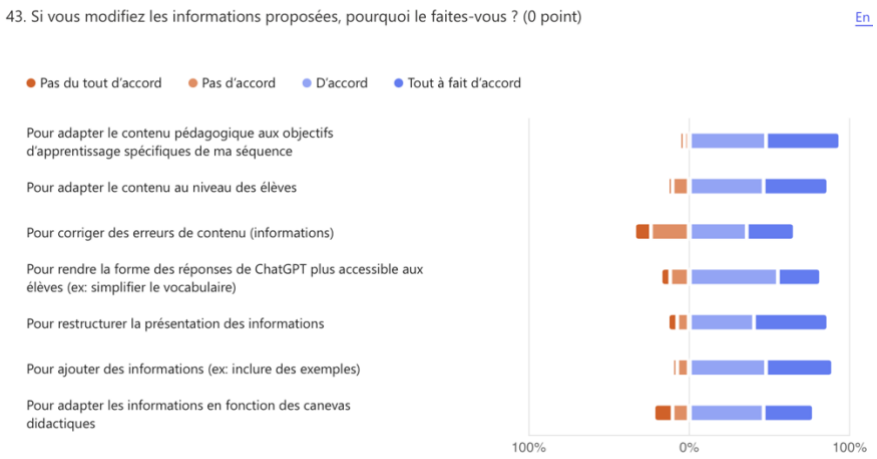


Figure 10: Survey – Question 43 – "If you modify the information given, what is the reason for doing so?" – Results.

Conclusion to RQ3-AThese numbers indicate that most teachers who use ChatGPT in their planning do not do so blindly. They review, adapt, and simplify the output to align it with learners' needs and contextual realities. The ability to critically assess and adapt AI-generated content is a key skill in responsible digital pedagogy, and it appears that many respondents have begun to develop this competence, even in the absence of formal training.

RQ2-F: *Do (future) language teachers use the premium version of ChatGPT?*
RQ3-B: *According to them, is there any difference in the quality of responses given to premium users and free users?*

The majority of respondents reported using only the free version of ChatGPT. When asked which version they currently use or have used:

- 77% reported using only the free version.
- 17% indicated that they had never used any version of ChatGPT.
- Only 5% had used both the free and premium versions.
- A single respondent (1%) reported using only the premium version.

Conclusion to RQ2-F: These figures confirm that premium usage remains extremely rare among language teachers in the sample. Although only a small number of respondents had experience with the premium version (6 in total), their qualitative feedback suggests that they did perceive advantages in terms of both pedagogy and performance. It is also noteworthy that three of these users are over the age of 30, which may suggest greater financial means to access paid tools.

Among the benefits mentioned were:

- "Pas de limites dans le nombre de prompts par jour »
- « Les réponses fournies sont nettement meilleures »
- « Possibilité de lui fournir un document »
- « Performance plus rapide et plus stable »

Conclusion to RQ3-B: Because the number of premium users was so limited, it is not possible to draw statistically valid conclusions. However, the consistency of these six responses supports the hypothesis that premium users experience improvements in both content quality and the overall user experience. Future research would be helpful to confirm these impressions.

5. Discussion

This discussion aims to interpret the findings of this study by drawing connections between the two complementary methodologies used: the analysis of the didactic sequences generated by ChatGPT, and the survey conducted among (future) language teachers in Liège. This section focuses on the parallels and contrasts that emerged between what the tool produced in practice and how it is perceived and used by educators in real educational contexts. This comparative approach helps highlight the broader implications of integrating ChatGPT into language education and provides a deeper understanding of its pedagogical potential and limitations.

The sequences created during the test phase showed that ChatGPT, particularly in its premium version, can generate structured and generally relevant content when provided with appropriate prompts (see RQ6). This aligns with research suggesting that AI can support creative lesson design and reduce teachers' workload (Ghafouri et al., 2024) (see RQ1). These observations were confirmed by the survey results, which showed that many teachers and trainees already use ChatGPT regularly for inspiration or to generate activities tailored to their students' needs (see RQ2). However, its effectiveness remained uneven across the different phases of the didactic sequence. While the tool respected general legal expectations and proposed realistic communicative tasks, it failed to include some essential pedagogical elements, such as metacognitive strategy exercises and logical coherence across the PFE model. The free version often showed clear limits, giving vague or messy answers and no authentic sources, unless the "search" function [27] was activated. These findings echo what participants in the survey expressed: teachers rarely use ChatGPT passively. Instead, they tend to critically evaluate and adapt its content before incorporating it into their teaching (see RQ3). This convergence suggests that, despite ChatGPT's potential, pedagogical expertise remains central to ensuring educational quality.

These shortcomings resonate strongly with the findings of the survey. Respondents regularly reported modifying ChatGPT's outputs before use in class, either to simplify the language, add missing didactic elements, or align tasks with learner needs. This critical engagement demonstrates that teachers do not use the tool passively but instead act as evaluators and adaptors of AI-generated material. This parallel between the test results and teacher behaviour

[27] The search option in ChatGPT lets it look things up on the internet to find the latest information.

suggests that while ChatGPT can assist with task generation, it requires human oversight to fulfill deeper pedagogical functions.

Another convergence between the two data sets relates to accessibility and training (see RQ2 and RQ4). The premium version of ChatGPT was found to produce more accurate sequences. Yet, most respondents indicated that they only had access to the free version. This disparity, between what the tool can technically deliver and what is practically available, highlights an important barrier in terms of digital equity. Teachers may lack the financial means or institutional support to access advanced tools, a point echoed in the literature (Rawas, 2024). Additionally, many survey participants expressed a clear need for targeted training, indicating that while digital curiosity exists, it is often unsupported by formal instruction or professional development.

Another relevant point of comparison lies in the way ChatGPT was used during the test and how teachers described using it in practice. In the testing phase, ChatGPT was not only asked to generate lessons but also to verify if the topic chosen ("Gap year") corresponded to the CEFR and curricula themes. This use (as a validation tool) was not widespread in the survey, but it represents a possible evolution of how AI might be used: not only to generate content but also to confirm its relevance. This suggests a potential added value for teachers who seek both inspiration and reassurance in their lesson planning.

Interestingly, the survey offered additional insights into personal and contextual factors (see RQ5). Differences were observed across age groups, experience levels, and status, with younger teachers generally more open to using ChatGPT. These variables were not directly explored through the tests but help contextualize how and why different teachers may or may not integrate ChatGPT into their practice.

Finally, both the tests and the survey reveal that the pedagogical effectiveness of ChatGPT depends on the human factor. Whether in guiding the tool with precise prompts or adapting its content, the teacher remains at the heart of the process. Teachers act as critical filters, ensuring that AI outputs meet the needs of their learners and align with institutional requirements. This reflects the broader pedagogical challenge of integrating AI in education: not simply adopting a new technology, but rethinking how it is used, evaluated, and taught.

In summary, both methodologies point to a common conclusion: ChatGPT is not a one-size-fits-all solution for language teaching, but a promising tool that can support creativity, reduce workload, and generate useful content, provided it is handled by trained and reflective teachers. As such, future training initiatives should focus not only on the technical use of the tool but also on developing teachers' evaluative and pedagogical skills, empowering them to use ChatGPT in ways that enhance, rather than replace, their professional expertise.

Conclusion

Synthesis

This dissertation aimed to explore whether ChatGPT can be used in English language education in the French-speaking Community of Belgium, with a particular focus on supporting (future) teachers. Through a combination of literature review, testing, and a survey of (future) teachers in Liège, the study provided insight into how this AI tool is used in practice, how it is perceived, what it can offer, where its limits lie, and what training needs to exist around its use.

The results show that ChatGPT (especially in its premium version) is capable of producing sequences that display a degree of coherence, structure, and alignment with the communicative and action-oriented approaches promoted in the CEFR and the WBF framework (RQ6). The tool demonstrated an ability to cover a wide range of language skills, integrate inductive grammar teaching, and reflect key pedagogical principles such as progression and learner engagement. However, this potential is conditional. While ChatGPT can generate structured sequences when prompted clearly and precisely, it struggles to autonomously integrate essential elements such as strategic instruction, metacognitive support, and context-specific differentiation. These shortcomings were particularly visible in sequences generated using the free version, which lacked internal coherence, included disjointed final tasks, and showed an inconsistent progression across phases, which ultimately makes it less pedagogically reliable and less suitable for classroom implementation.

The survey results confirmed these observations from the test phase. Most respondents reported using ChatGPT in their professional or academic practices (RQ2), though typically in a selective and critical manner. Rather than applying AI-generated materials directly, teachers described adapting, supplementing, or modifying them—often to ensure they met the specific needs of learners or to align with legal and pedagogical guidelines (RQ3). This confirms that

while ChatGPT can reduce workload and stimulate creativity (RQ1), it is not perceived as a substitute for teacher expertise, but rather as a tool to be guided and reviewed. The study also examined the influence of contextual and personal factors (RQ5), finding only limited variation in attitudes across demographics such as age or teaching experience.

One of the most striking findings regards the question of training (RQ4). Although many teachers expressed interest in using ChatGPT, the majority had not received formal instruction on how to do so effectively. The lack of structured training was identified as a major barrier to confident integration. This suggests an urgent need to incorporate more AI literacy into both initial teacher education and continuing professional development programs.

In light of these findings, it becomes clear that ChatGPT holds genuine potential to assist language teachers in designing didactic sequences, especially when paired with models such as the PFE, which can provide the pedagogical structure that the tool itself lacks. However, to fully harness this potential, educational institutions must support teachers not only with access to the most effective versions of the tool but also with training focused on critical use, prompt design, and pedagogical adaptation.

Finally, this dissertation contributes to the broader conversation about AI in education by highlighting the need for thoughtful, teacher-led integration. ChatGPT is not a replacement for the teacher, who remains the central figure responsible for delivering the course, but a partner that requires human expertise to be used meaningfully. If approached critically and ethically, it can be a powerful ally in fostering innovation, reducing workload, and enhancing the quality of language education in a rapidly evolving digital landscape.

Usage perspectives

As mentioned in the literature review, several experts have warned that AI tools like ChatGPT can sometimes reproduce or reinforce stereotypes (Thorne, 2024). Since these tools are trained on massive amounts of online data, they may reflect the biases that exist in society. To explore this idea further, I tested ChatGPT's image generation feature with a few prompts[28]. When I asked it to create an image of a "happy family", the result showed only white people. In another test, I asked for an image where one person was cooking and another was watching TV. The

[28] See Appendix 8

image showed a woman in the kitchen and a man on the sofa, which clearly reinforces a traditional gender stereotype. I also asked for an image of a basketball player, and the person generated was Black. Curious, I asked ChatGPT why the player was Black. The tool replied that it "sometimes draws from common visual patterns or biases in media representations", and that these patterns can unfortunately reinforce stereotypes. While ChatGPT is capable of adjusting its responses when asked, for example, by generating more diverse or balanced images, these tests confirm what researchers like Thorne (2024) have argued: AI tools reflect the content they are trained on, and without clear, thoughtful instructions, they can unintentionally reproduce existing societal biases, much like those already present in educational manuals. This shows just how important it is for teachers to use these tools critically and to stay alert to the representations they generate in the classroom. AI can be a helpful resource, but only if it is guided in ways that promote inclusion and avoid reinforcing outdated or unfair stereotypes.

In addition, during the test, one of the main weaknesses observed in the didactic sequence produced by ChatGPT was the limited number of exercises, particularly in the Fixation phase. The tool provided only a few activities to reinforce vocabulary and grammar, and some lessons were overly focused on preparing the final task rather than consolidating language structures. However, it is important to note that this limitation is not due to the tool's incapacity, but rather to how it was initially prompted. When asked more specifically to generate additional or varied exercises[29], ChatGPT was able to do so quickly and effectively. This suggests that while the default output may lack pedagogical depth in some areas, the tool remains flexible and responsive when guided with clear, detailed instructions. Therefore, with better prompting, it is possible to improve the quantity and diversity of activities, making the Fixation phase more complete and aligned with good teaching practices.

In the literature review, some researchers pointed out that ChatGPT may lack emotional sensitivity or the ability to provide empathetic feedback to students (Nizzolino, 2024). This limitation is often mentioned as a major drawback of AI in education, especially when it comes to building learner confidence or responding to emotional needs. To explore this concern, I decided to test ChatGPT's ability to respond with empathy in a realistic classroom scenario[30]. I asked the tool how it would give feedback to a student who had made mistakes, and how it

[29] See Appendix 9
[30] See Appendix 10

could reassure or support that learner. The response was surprisingly thoughtful: ChatGPT offered constructive feedback, highlighted the student's efforts, encouraged progress, and even suggested phrases a teacher could use to boost confidence. While it is clear that ChatGPT cannot truly understand emotions, this test showed that it can simulate emotional support when prompted correctly. This suggests that, although limited, the tool can still be used to model or support positive and encouraging teacher language, especially when guided by a user who is aware of the importance of emotional sensitivity in the learning process.

Research perspectives

While this study focused on evaluating the coherence of a didactic sequence produced by ChatGPT and analysing the attitudes of (future) language teachers, several promising avenues for further research and pedagogical innovation remain to be explored.

One key area is the potential of ChatGPT to support differentiation in the language classroom. Because it can quickly adjust content to fit students' levels, needs, or interests, ChatGPT can help teachers make activities for classes with different abilities or give extra support to students who need it. Future research could examine how teachers can use AI to create parallel tasks within the same lesson, adapted to different profiles, while still meeting curricular objectives.

Another key area is using ChatGPT for speaking practice. This study focused mainly on creating sequences, but ChatGPT's ability to hold conversations could help students improve their speaking skills through role-plays, simulations, or pronunciation help. Exploring how it could prepare students for real-life situations, especially interactive ones, would be a useful step in using AI in education.

In addition, ChatGPT's potential for providing personalized evaluation and feedback deserves attention. Although this aspect was not explored in the current study, the tool can offer instant responses, corrections, and suggestions, which could help teachers in both formative and summative assessment. Future research might investigate the reliability and relevance of such feedback, especially in terms of alignment with official assessment documents like the CEFR, frameworks and curricula.

The issue of academic integrity and detection tools also remains central. As showed in the literature and confirmed by respondents in the survey, many teachers worry about cheating or over-reliance on AI-generated content. Exploring the reliability of detection tools such as

ZeroGPT, Turnitin, or GPTZero, and how teachers use them (or do not), would help clarify how AI can be responsibly managed in an educational context.

It would also be relevant to compare the performance of different generative AI models. ChatGPT is currently one of the most well-known tools, but others, such as Gemini, Claude, and DeepSeek, are rapidly evolving. A comparative study using the same methodology as this disssertation, but with a model like DeepSeek, could provide valuable insights into how various platforms differ in terms of pedagogical coherence, content quality, and usability for language teachers.

Moreover, this paper focused exclusively on the perceptions and practices of teachers and future teachers. A similar study investigating the perspectives of students (how they experience ChatGPT in the classroom, how they use it independently, and how they think it could support their learning) would offer valuable complementary insights. Understanding students' attitudes and expectations could help ensure that AI integration supports not only teaching quality but also learner engagement and autonomy.

It is also important to note that this study tested ChatGPT's capabilities within the framework of the PFE model. However, other didactic approaches exist, such as the problem-based model. Future research could explore how ChatGPT performs within these alternative models, to evaluate its adaptability and relevance across different pedagogical approaches.

Another relevant area for future research is the environmental impact of generative AI. Tools like ChatGPT rely on high computational power, which consumes energy and contributes to carbon emissions. While these tools offer real pedagogical benefits, their use raises important sustainability questions. Future studies could explore how often and in what contexts AI should be used in education to ensure a balance between innovation and environmental responsibility.

Finally, as tools like ChatGPT continue to evolve, their changing abilities highlight the need for regular updates in research and teacher training. Since ChatGPT's responses are constantly improving, any evaluation should be seen as a snapshot of a specific moment. This means teachers (both future and current) need ongoing support to follow these changes. Future research could look at how teacher education programmes help teachers use AI in a critical, creative, and ethical way, so they feel confident and prepared to include it in their lessons.

In short, this dissertation represents only a first step in understanding how generative AI can contribute to language education. The perspectives outlined here show that much remains to be explored, especially in ways that promote learner engagement, teacher autonomy, and pedagogical quality.

Bibliography

Ahn, J., Lee, J., & Son, M. (2024). ChatGPT in ELT: disruptor? Or well-trained teaching assistant?. *ELT Journal*, *78*(3), 345–355. https://doi.org/10.1093/elt/ccae017

Al-khresheh, M. H. (2024). Bridging technology and pedagogy from a global lens: Teachers' perspectives on integrating ChatGPT in English language teaching. *Computers and Education. Artificial Intelligence*, *6*, 1-12. https://doi.org/10.1016/j.caeai.2024.100218

Annamalai, N. (2024). Factors affecting English language high school teachers switching intention to ChatGPT: a Push-Pull-Mooring theory perspective. *Interactive Learning Environments*, *33*(2), 1-18. https://doi.org/10.1080/10494820.2024.2371928

Baskara, F. R. (2023). AI-Driven Dynamics: ChatGPT Transforming ELT Teacher-Student Interactions. *Lensa: Kajian Kebahasaan, Kesusastraan, Dan Budaya*, *13*(2), 261–275. https://doi.org/10.26714/lensa.13.2.2023.261-275

Cambridge Dictionary (n.d). *Interoperability.* https://dictionary.cambridge.org/dictionary/english/interoperability

Chan, C. K. Y., & Lee, K. K. W. (2023). The AI generation gap: Are Gen Z students more interested in adopting generative AI such as ChatGPT in teaching and learning than their Gen X and millennial generation teachers?. *Smart Learning Environments*, *10*(1), 60–23. https://doi.org/10.1186/s40561-023-00269-3

Cihon, P. (2019). Standards for AI governance: international standards to enable global coordination in AI research & development. *Future of Humanity Institute*, *40*(3), 340-342.

Council of Europe (2001). *Common European Framework of Reference for Languages: learning, teaching, Assessment*. Cambridge University Press

Council of Europe. (2018). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment – Companion Volume with New Descriptors*. Council of Europe Publishing.

Dalalah, D., & Dalalah, O. M. A. (2023). The false positives and false negatives of generative AI detection tools in education and academic research: The case of ChatGPT. *The International Journal of Management Education*, *21*(2), 1-13. https://doi.org/10.1016/j.ijme.2023.100822

Digital Wallonia (n.d). *AI for Edu*. www.digitalwallonia.be

Du, J., & Alm, A. (2024). The Impact of ChatGPT on English for Academic Purposes (EAP) Students' Language Learning Experience: A Self-Determination Theory Perspective. *Education Sciences*, *14*(726), 1-17. https://doi.org/10.3390/educsci14070726

EduLAB (n.d). *Parcours référents IA.* https://edu-lab.be

Elkhatat, A. M., Elsaid, K., & Almeer, S. (2023). Evaluating the efficacy of AI content detection tools in differentiating between human and AI-generated text. *International Journal for Educational Integrity*, *19*(1), 1–16. https://doi.org/10.1007/s40979-023-00140-5

Encyclopedia (n.d). *Intelligent Tutoring System (ITS)*. https://encyclopedia.pub/entry/28717

Engati (n.d). *Rule-based systems.* https://www.engati.com/glossary/rule-based-system

Faculté des Sciences Appliquées. (2024). *La perspective des étudiants sur l'utilisation de l'IA dans l'enseignement en FSA*. Université de Liège.

Fédération de l'Enseignement Secondaire Catholique (FESeC). (2018). *Programme pour les 2e et 3e degrés – Langues modernes I, II, III.* Fédération de l'Enseignement Secondaire catholique.

Fédération de l'Enseignement Secondaire Catholique (FESeC). (2017). *Programme des humanités professionnelles et techniques : 2e et 3e degrés – Professionnel et Technique de qualification – Formation générale commune – Langues modernes.* Fédération de l'Enseignement Secondaire catholique.

Fédération Wallonie-Bruxelles. *Enseignement*. http://enseignement.be/index.php

García-López, I. M., González, C. S. G., Ramírez-Montoya, M. S., & Molina-Espinosa, J. M. (2025). Challenges of implementing ChatGPT on education: Systematic literature review. *International Journal of Educational Research Open*, *8*(100401), 1-13. https://doi.org/10.1016/j.ijedro.2024.100401

Ghafar, Z. N. (2023). ChatGPT: a new tool to improve teaching and evaluation of second and foreign languages a review of ChatGPT: the future of education. *International journal of applied research and sustainable sciences*, *1*(2), 73-86.

Ghafouri, M., Hassaskhah, J., & Mahdavi-Zafarghandi, A. (2024). From virtual assistant to writing mentor: Exploring the impact of a ChatGPT-based writing instruction protocol on EFL teachers' self-efficacy and learners' writing skill. *Language Teaching Research: LTR*, 1-23. https://doi.org/10.1177/13621688241239764

Guo, K., & Wang, D. (2024). To resist it or to embrace it? Examining ChatGPT's potential to support teacher feedback in EFL writing. *Education and Information Technologies*, *29*(7), 8435–8463. https://doi.org/10.1007/s10639-023-12146-0

Guo, Y., & Lee, D. (2023). Leveraging ChatGPT for Enhancing Critical Thinking Skills. *Journal of Chemical Education*, *100*(12), 4876–4883. https://doi.org/10.1021/acs.jchemed.3c00505

HerftEducator (n.d). *A teacher's prompt guide to ChatGPT aligned with 'What Works Best'.* https://www.herfteducator.com

IBM (n.d). *What is generative AI?* https://research.ibm.com/blog/what-is-generative-AI

IBM (n.d). *Algorithmic bias.* https://www.ibm.com/think/topics/algorithmic-bias

Institut Interréseaux de la Formation Professionnelle Continue (IFPC) (n.d). *Formations.* https://ifpc-fwb.be/v5/default.asp

Investopedia (n.d). *Turing tests.* https://www.investopedia.com/terms/t/turing-test.asp

Juan, W., Qing, S., Yunus, M. M., & M. Rafiq, K. R. (2023). Integrating ChatGPT into English Language Teaching and Learning: Strengths and Weaknesses. *International Journal of Academic Research in Business & Social Sciences*, *13*(12), 4055-4068. https://doi.org/10.6007/IJARBSS/v13-i12/20251

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for Language Teaching and Learning. *RELC Journal*, *54*(2), 537–550. https://doi.org/10.1177/00336882231162868

Koraishi, O. (2023). Teaching English in the age of AI: Embracing ChatGPT to optimize EFL materials and assessment. *Language Education and Technology*, *3*(1), 55-72.

Kostka, I., & Toncelli, R. (2023). Exploring Applications of ChatGPT to English Language Teaching: Opportunities, Challenges, and Recommendations. *TESL-EJ*, *27*(3), 1-19. https://doi.org/10.55593/ej.27107int

Malik, M. A., & Amjad, A. I. (2025). AI vs AI: How effective are Turnitin, ZeroGPT, GPTZero, and Writer AI in detecting text generated by ChatGPT, Perplexity, and Gemini?. *Journal of Applied Learning and Teaching*, *8*(1), 91-101.

Mitsloan (n.d). *Machine learning explained*. https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained

Nabizadeh Rafsanjani, H., & Nabizadeh, A. H. (2023). Towards human-centered artificial intelligence (AI) in architecture, engineering, and construction (AEC) industry. *Computers in Human Behavior Reports*, *11*, 1-9. https://doi.org/10.1016/j.chbr.2023.100319

Nizzolino, S. (2024). Artificial Intelligence in Language Teaching: Using ChatGPT to assist teachers of English as a Foreign Language. *Form@re*, *24*(1), 242-261. https://doi.org/10.36253/form-15681

Omohundro, S. M. (2018). The basic AI drives. *Artificial intelligence safety and security, Chapman and Hall/CRC*, 47-55.

OpenAI (n.d). *Index ChatGPT*. https://openai.com/index/chatgpt/

Paganelli, T. (2024, February 14). *Choix et intégration des outils numériques à ma pratique pédagogique en langues modernes* [Presentation]. University of Liège, Belgium.

Pradeep, A., Satmuratov, A., Yeshbayev, I., Khasan, O., Iqboljon, M., & Daniyor, A. (2023, August). The significance of artificial intelligence in contemporary cinema. In *2023 Second International Conference on Trends in Electrical, Electronics, and Computer Engineering (TEECCON)*, 1-6.

Rawas, S. (2024). ChatGPT: Empowering lifelong learning in the digital age of higher education. *Education and Information Technologies*, *29*(6), 6895-6908.

Sarrion, E. (2023). *What Is ChatGPT? Exploring the power of ChatGPT: Applications, techniques, and implications*, Apress.

Service général de l'Enseignement organisé par la Fédération Wallonie-Bruxelles. (2017). *Compétences terminales et savoirs requis à l'issue des humanités générales et technologiques– Langues modernes.*

Service général de l'Enseignement organisé par la Fédération Wallonie-Bruxelles (2020). *Programme d'études des langues modernes. Humanités générales et technologiques : 2e et 3e degrés.*

Service général de l'Enseignement organisé par la Fédération Wallonie-Bruxelles (2020). *Programme d'études des langues modernes. Humanités professionnelles et techniques : 2e et 3e degrés.*

Service général de l'Enseignement organisé par la Fédération Wallonie-Bruxelles et Conseils des Pouvoirs Organisateurs de l'Enseignement Officiel Neutre Subventionné (2018). *Langues modernes : Outil d'accompagnement du référentiel « Compétences terminales et savoirs requis à l'issus des humanités générales et technologiques ».* 2ème degré de l'enseignement général et technique de transition. LM 1. 4p/s.

Simons, G. (2023). *AESS0329-1 : Didactique des langues modernes. Partim I.* Department of Modern Languages, University de Liège.

Simons, G. (2018). De la place des genres textuels dans les familles de tâches en langues modernes. In M. Jacquin, G. Simons, & D. Delbrassine (Eds.), *Les genres textuels en langues étrangères : entre théorie et pratique* (37–70). Peter Lang Group.

Stutz, P., Elixhauser, M., Grubinger-Preiner, J., Linner, V., Reibersdorfer-Adelsberger, E., Traun, C., Wallentin, G., Wöhs, K., & Zuberbühler, T. (2023). Ch(e)atGPT? An Anecdotal Approach Addressing the Impact of ChatGPT on Teaching and Learning GIScience. *GI Forum*, *11*(1), 1-8. https://doi.org/10.1553/giscience2023_01_s140

Suriano, R., Plebe, A., Acciai, A., & Fabio, R. A. (2025). Student interaction with ChatGPT can promote complex critical thinking skills. *Learning and Instruction*, *95*, 1-8. https://doi.org/10.1016/j.learninstruc.2024.102011

Tahiru, F. (2021). AI in education: A systematic literature review. *Journal of Cases on Information Technology (JCIT)*, *23*(1), 1-20.

Techtarget (n.d). *Natural Language Processing (NLP)*. https://www.techtarget.com/searchenterpriseai/definition/natural-language-processing-NLP

Thorne, S. (2025, February). *Interfacing langage learning in the wild with classroom instruction* [Presentation]. University of Liège, Belgium.

Thorne, S. L. (2024). Generative artificial intelligence, co-evolution, and language education. *Modern Language Journal*, *108*, 567-572. https://doi.org/10.1111/modl.12932

Thorne, S. (2025, February). *Generative AI for L2 learning: analysis of ChatGPT's discourse* [Presentation]. University of Liège, Belgium.

Toosi, A., Bottino, A. G., Saboury, B., Siegel, E., & Rahmim, A. (2021). A brief history of AI: how to prevent another winter (a critical review). *PET clinics*, *16*(4), 449-469.

Trust, T., Whalen, J., & Mouza, C. (2023). Editorial: ChatGPT: challenges, opportunities, and implications for teacher education. *Contemporary Issues in Technology and Teacher Education*, *23* (1), 1–23.

Ulla, M. B., Perales, W. F., & Busbus, S. O. (2023). To generate or stop generating response: Exploring EFL teachers' perspectives on ChatGPT in English language teaching in Thailand. *Learning*, *9*(2), 168–182. https://doi.org/10.1080/23735082.2023.2257252

UNESCO (n.d). *Generative AI: UNESCO study reveals alarming evidence of regressive gender stereotypes*. https://www.unesco.org/en/articles/generative-ai-unesco-study-reveals-alarming-evidence-regressive-gender-stereotypes

Urazbayeva, G., Kussainova, R., Aibergen, A., Kaliyeva, A., & Kantayeva, G. (2024). Innovation Off the Bat: Bridging the ChatGPT Gap in Digital Competence among English as a Foreign Language Teachers. *Education Sciences*, *14*(946), 1-15. https://doi.org/10.3390/educsci14090946

van den Berg, G., & du Plessis, E. (2023). ChatGPT and Generative AI: Possibilities for Its Contribution to Lesson Planning, Critical Thinking and Openness in Teacher Education. *Education Sciences*, *13*(10), 1-12. https://doi.org/10.3390/educsci13100998

Working group (ISLV). (2024). *Intelligences artificielles génératives et enseignement/apprentissage des langues étrangères*. University of Liège.

Yuan, Y., Li, H., & Sawaengdist, A. (2024). The impact of ChatGPT on learners in English academic writing: opportunities and challenges in education. *Language Learning in Higher Education*, *14*(1), 41–56. https://doi.org/10.1515/cercles-2023-0006

Zeevy Solovey, O. (2024). Comparing peer, ChatGPT, and teacher corrective feedback in EFL writing: Students' perceptions and preferences. *Technology in Language Teaching and Learning*, *6*(3), 1-23. https://doi.org/10.29140/tltl.v6n3.1482