



Université Mohamed Khider de Biskra  
Faculté des Lettres et des Langues  
Département de langue et littérature françaises  
Filière de français

# MÉMOIRE DE MASTER

Option : Didactique langues cultures

---

BENAMMAR Ibtissem

Présenté par:

Dre REFRAFI Soraya

## Évaluer la compréhension de l'oral en FLE par l'IA : étude de l'apport de ChatGPT auprès d'étudiants de 3e année licence à l'université de Biskra-

---

Jury

Année universitaire : 2024/2025

## Remerciements

Au nom d'Allah, Le Tout-Puissant, Le Miséricordieux, qui m'a guidée sur le droit chemin, m'a accordé la santé, la volonté et la force d'accomplir ce travail.

Je tiens à exprimer ma profonde gratitude à mon enseignante et encadrante, Madame Soraya REFRAFI, pour son accompagnement attentif et ses conseils éclairés tout au long de la réalisation de ce mémoire. Sa bienveillance, sa rigueur et ses encouragements constants m'ont permis de progresser, de croire en mes capacités et de donner le meilleur de moi-même. Grâce à elle, j'ai franchi des étapes importantes dans mon parcours académique. Merci infiniment pour votre engagement, votre patience et la lumière que vous avez apportée à mon chemin.

Je remercie également l'ensemble de mes professeurs de l'Université de Biskra, en particulier les membres du jury, pour le temps qu'ils ont consacré à la lecture et à l'évaluation de ce travail.

J'adresse un remerciement très spécial à ma famille, pour son soutien inconditionnel, sa patience et ses sacrifices, sans lesquels cette aventure n'aurait pas été possible.

Enfin, je remercie toutes les personnes qui m'ont aidée, de près ou de loin, dans l'élaboration de ce mémoire.

## *Dédicace*

*Je dédie ce modeste travail à mes chers parents.*

*À mon père, Lazhar , pilier de force et de sagesse, merci pour ton soutien constant, ta confiance inébranlable et tes conseils toujours empreints de lucidité, même dans les moments les plus injustes que j'ai traversés. Tu as cru en moi lorsque moi-même j'en doutais.*

*À ma mère, Samia , source infinie d'amour, de patience et de prières silencieuses, merci d'avoir essuyé mes larmes et soutenu mon cœur dans les instants les plus douloureux. Ton affection m'a portée là où je suis aujourd'hui.*

*À mes frères, Charaf Eddine, Moutez Bellah, Mohammed Abdou et Mouatassam Bellah, pour votre présence silencieuse mais si forte, vos mots simples, votre fraternité sincère. Vous êtes des piliers dans ma vie.*

*À ma sœur unique, Khawla Ben Ammar, épouse dévouée et mère admirable, merci pour ton amour, ta sagesse et ta force. Tu es un repère inestimable dans ma vie. À tes enfants, Bara, Ishek et Mohammed Anès.*

*Et à ma chère amie Fatima Rahma, dont l'amitié sincère a été une lumière discrète mais essentielle.*

# TABLE DES MATIERES

<b>Introduction générale.....</b>	<b>7</b>
<b>Chapitre I : concepts fondamentaux .....</b>	<b>11</b>
<b>Introduction.....</b>	<b>12</b>
<b>1. Définition et rôle de l'évaluation en didactique des langues.....</b>	<b>12</b>
1.1. L'évaluation formative et sommative .....	13
1.2. Les objectifs de l'évaluation dans l'enseignement des langues.....	13
1.3. Le rôle de l'évaluation dans le FLE .....	14
<b>2 L'intelligence artificielle : définition, types et applications en didactique des langues</b> .....	<b>15</b>
2.1 Définition de l'intelligence artificielle.....	15
2.2 Les principaux types d'intelligence artificielle.....	16
2.3 Types d'IA utilisés en didactique des langues .....	17
2.3. L'introduction de l'intelligence artificielle dans l'éducation.....	18
<b>3. Les principes de l'évaluation dans l'enseignement des langues .....</b>	<b>21</b>
3.1. Objectifs de l'évaluation dans l'enseignement des langues.....	21
3.2. Les principes fondamentaux d'une évaluation de qualité.....	22
3.3. Vers une évaluation formative, inclusive et réflexive .....	23
<b>4. Innovations : l'apport de l'intelligence artificielle à l'évaluation en FLE.....</b>	<b>24</b>
4.1. Les apports de l'IA à l'évaluation linguistique .....	24
4.1.3. Reconnaissance vocale pour évaluer l'oralité.....	25
4.2. Vers une évaluation plus personnalisée et adaptative .....	26
4.3 Limites de l'évaluation par l'IA en FLE .....	26
3.4 Vers une complémentarité entre l'enseignant et l'intelligence artificielle.....	27
<b>Conclusion .....</b>	<b>27</b>

<b>Chapitre II : Étude expérimentale sur l'évaluation de la compréhension orale assistée ou non par IA .....</b>	<b>29</b>
<b>Introduction.....</b>	<b>30</b>
<b>1. Cadre méthodologique .....</b>	<b>31</b>
1.1 Objectifs de l'expérimentation.....	31
1.2. Participants.....	31
1.2. Matériel utilisé .....	31
1.3. Déroulement général .....	32
1.4. Outil complémentaire d'observation : la grille.....	32
<b>2. Description des modalités expérimentales.....</b>	<b>34</b>
2.1. Protocole pour le groupe témoin : évaluation manuelle classique.....	34
2.2. Protocole pour le groupe expérimental : évaluation assistée par IA.....	35
2.3. Comparaison des deux modalités.....	36
<b>3.Déroulement de l'expérimentation : Groupe témoin (évaluation classique).....</b>	<b>36</b>
3.1. Phase de pré-visionnage.....	36
3.2. Visionnage de la capsule vidéo .....	37
3.3. Passation du questionnaire de compréhension.....	37
3.4. Traitement des réponses et évaluation .....	37
<b>4. Protocole pour le groupe expérimental : évaluation assistée par intelligence artificielle.....</b>	<b>38</b>
4.1. Conditions matérielles et pédagogiques.....	38
4.2. Déroulement de la séance .....	38
4.3. Contenu du questionnaire .....	39
4.4. Procédure d'évaluation automatisée .....	39
4.5. Délai et vérification.....	39
4.6. Retour aux étudiants .....	39
<b>1. Comparaison des deux modalités d'évaluation .....</b>	<b>40</b>
5.1. Critères comparatifs retenus .....	40

5.2. Résultats .....	40
6.2. Limites de l'expérimentation .....	47
6.3. Conditions nécessaires à une intégration réussie de l'IA.....	48
<b>Conclusion .....</b>	<b>49</b>
<b>Conclusion générale .....</b>	<b>50</b>
<b>Références Bibliographiques .....</b>	<b>54</b>
<b>ANNEXES</b>	<b>58</b>
<b>Annexe 1 – Questionnaire de compréhension orale.....</b>	<b>59</b>
<b>Annexe 2 – Grille d'évaluation : Groupe Témoin.....</b>	<b>60</b>
<b>Annexe 3 – Grille d'évaluation : Groupe Expérimental .....</b>	<b>61</b>
<b>Annexe 3 Exemple de Réponse d'étudiant – Groupe Expérimental avec feedback ...</b>	<b>62</b>
<b>Annexe 4 : Description de la vidéo utilisée .....</b>	<b>63</b>

## **INTRODUCTION GENERALE**

L'enseignement et l'évaluation des langues étrangères, en particulier le Français Langue Étrangère (FLE), ont vu leurs pratiques évoluer avec l'introduction de technologies de plus en plus sophistiquées. Parmi ces technologies, l'intelligence artificielle (IA) a pris une place importante, notamment dans la personnalisation de l'apprentissage et l'automatisation des évaluations. L'émergence des outils numériques et des IA dans l'éducation, à la fois comme support pédagogique et comme outil d'évaluation, soulève des questions sur la pertinence de leur usage dans le domaine de l'enseignement des langues.

L'un des enjeux majeurs réside dans la capacité de ces technologies à remplacer ou compléter l'évaluation traditionnelle réalisée par un enseignant. Ce mémoire s'intéresse particulièrement à l'utilisation de l'IA, et plus spécifiquement de ChatGPT, dans l'évaluation de la compréhension orale en FLE. L'objectif est de comparer l'efficacité et la rapidité des retours fournis par une évaluation automatisée avec celle d'une évaluation humaine traditionnelle. Cette réflexion se situe dans un contexte où l'IA pourrait potentiellement transformer les pratiques pédagogiques et offrir des réponses aux défis contemporains liés à l'enseignement à distance et à la personnalisation de l'évaluation.

L'objectif principal de ce travail est de déterminer si l'évaluation par IA, en l'occurrence via ChatGPT, peut être aussi efficace qu'une évaluation humaine dans un cadre académique d'enseignement du FLE. L'expérimentation menée se concentre sur l'évaluation de la compréhension orale, en comparant les performances des étudiants lorsqu'ils reçoivent un feedback automatisé par IA et un feedback humain.

Plus précisément, les objectifs sont les suivants :

- Évaluer la rapidité et la qualité des retours fournis par ChatGPT en comparaison avec un enseignant humain.
- Analyser la capacité de l'IA à fournir des feedbacks détaillés et constructifs pour améliorer les compétences des étudiants.
- Identifier les avantages et les limites de l'IA dans l'évaluation des compétences linguistiques par rapport aux méthodes d'évaluation classiques.

La question centrale de ce travail est : L'intelligence artificielle peut-elle remplacer ou compléter efficacement l'évaluation humaine dans le cadre de l'enseignement du FLE ? Les questions spécifiques sont :



- En quoi l'IA, comme ChatGPT, peut-elle être un outil pertinent pour l'évaluation des compétences linguistiques des étudiants en FLE ?
- Quel impact l'évaluation par IA a-t-elle sur la motivation des étudiants et leur engagement dans le processus d'apprentissage ?
- L'évaluation automatisée par IA est-elle aussi juste et fiable qu'une évaluation réalisée par un enseignant ?

Le cadre théorique de ce travail repose sur plusieurs axes, notamment l'évaluation en pédagogie, les théories de l'apprentissage des langues, et les applications de l'IA dans l'éducation. L'évaluation en didactique des langues a toujours été un sujet central, et des chercheurs comme Ducrot (2004) ont montré l'importance du feedback dans le processus d'apprentissage. Ce feedback, considéré comme un vecteur de motivation, permet à l'apprenant de corriger ses erreurs et d'améliorer ses compétences linguistiques.

Dans le domaine de l'IA, des travaux comme ceux de Mialaret (2018) ont souligné l'impact potentiel des outils d'IA sur l'automatisation des évaluations, tout en soulevant la question de leur fiabilité dans le cadre de l'enseignement des langues. Collins & Hull (2017) ont abordé l'importance de l'adaptabilité des systèmes d'IA pour offrir une évaluation personnalisée et justifiée.

Cependant, l'introduction de l'IA dans l'évaluation n'est pas sans critique. Lemoine (2019) souligne que l'IA, bien qu'efficace pour générer des réponses, ne peut pas se substituer à l'intelligence humaine, notamment lorsqu'il s'agit d'évaluer des éléments complexes tels que la compréhension contextuelle ou la subjectivité d'un discours.

La question de l'efficacité de l'évaluation par IA est d'une importance croissante, car l'enseignement des langues subit une digitalisation accélérée, notamment avec l'essor des technologies et des plateformes d'apprentissage en ligne. De plus, la crise sanitaire liée à la pandémie de COVID-19 a accéléré l'adoption des outils numériques, faisant de l'IA un acteur essentiel dans l'optimisation des pratiques pédagogiques.

Cette recherche est d'autant plus pertinente qu'elle s'inscrit dans un contexte où l'IA pourrait répondre à des besoins spécifiques dans l'enseignement des langues, tels que la rapidité d'évaluation, la personnalisation du feedback, et la réduction des biais humains. En outre, elle permet de mettre en lumière les limites et les possibilités d'un outil comme ChatGPT dans l'accompagnement des étudiants en FLE.

Ce mémoire se structure en deux chapitres principaux :

- **Le Chapitre I – Concepts fondamentaux** : Cette première partie examine les concepts clés liés à l'évaluation en pédagogie, les spécificités de l'évaluation en FLE, ainsi que l'application de l'IA dans l'enseignement des langues. Elle explore également les différentes théories de l'apprentissage et leur lien avec les nouvelles technologies.
- **Le Chapitre II – Étude expérimentale sur l'évaluation de la compréhension orale assistée ou non par IA** : Cette seconde partie présente l'expérimentation menée auprès des étudiants de FLE, les méthodologies utilisées, les résultats obtenus, et une discussion approfondie sur les implications pédagogiques de l'évaluation par IA. Elle compare également les retours donnés par l'IA et ceux fournis par un enseignant humain.

Ce mémoire vise à apporter des éléments de réponse aux questions soulevées par l'introduction de l'IA dans l'évaluation des compétences linguistiques, en particulier dans le domaine du FLE. En explorant les avantages et les limitations de l'évaluation automatisée par IA, cette étude contribuera à la réflexion sur l'avenir de l'enseignement des langues et la place croissante de l'intelligence artificielle dans l'éducation.

## **CHAPITRE I : CONCEPTS FONDAMENTAUX**

## **Introduction**

L'évaluation joue un rôle clé dans l'enseignement des langues, notamment du français langue étrangère (FLE). Elle permet non seulement de mesurer les progrès des apprenants, mais aussi d'ajuster les pratiques pédagogiques et d'orienter les futures stratégies d'enseignement. Depuis plusieurs années, de nombreux modèles ont été proposés pour éclairer les différentes formes d'évaluation, qu'elle soit formative, sommative ou diagnostique. L'objectif reste de comprendre les attentes des apprenants et de répondre aux spécificités du contexte éducatif.

Avec les avancées technologiques, l'évaluation des compétences en langues a évolué, notamment avec l'introduction de l'intelligence artificielle (IA). Cette technologie, qui regroupe un ensemble d'outils permettant de simuler des processus cognitifs humains, trouve de plus en plus sa place dans les pratiques pédagogiques. Appliquée à l'évaluation, l'IA permet d'automatiser certains aspects du processus, rendant les corrections plus rapides et personnalisées, en particulier grâce au traitement du langage naturel. Les outils basés sur l'IA offrent ainsi de nouvelles possibilités dans l'analyse des compétences langagières.

Cependant, l'utilisation de l'IA dans l'évaluation soulève diverses interrogations sur ses avantages réels, mais aussi ses limites. Alors que certains chercheurs affirment que l'IA rend l'évaluation plus rapide et plus précise (Gérard et Lemoine, 2019), d'autres soulignent les défis liés à la prise en compte des nuances culturelles et contextuelles des contenus évalués (Zammit et al., 2021). En particulier, l'évaluation de la compréhension orale en FLE demeure un domaine complexe, en raison de la nécessité de tenir compte de facteurs subjectifs, comme la prosodie, l'intonation et le contexte.

Ce chapitre vise à présenter les fondements théoriques de l'évaluation dans l'enseignement des langues, en analysant les modèles qui ont influencé les pratiques actuelles. Il s'agira également de discuter des implications de l'intégration de l'IA dans ce domaine, en se concentrant sur ses applications, ses avantages et ses limites. Cette réflexion servira de base pour le chapitre suivant, dans lequel sera mise en place une expérimentation concrète pour tester l'efficacité de l'IA dans l'évaluation de la compréhension orale en FLE.

### **1. Définition et rôle de l'évaluation en didactique des langues**

L'évaluation, dans le cadre de la didactique des langues, ne se limite pas à une simple mesure des connaissances linguistiques des apprenants. Elle constitue un outil central

dans la mise en œuvre des pratiques pédagogiques, permettant non seulement d'apprécier les acquis mais aussi d'adapter les stratégies d'enseignement pour répondre aux besoins individuels des apprenants. Selon certains didacticiens, « l'évaluation est un dispositif qui permet à la fois de rendre compte des progrès de l'apprenant et de guider l'enseignement vers de nouveaux objectifs » (Rivière, 2017).

### **1.1. L'évaluation formative et sommative**

Dans la didactique des langues, deux types principaux d'évaluation sont régulièrement utilisés : l'évaluation formative et l'évaluation sommative.

- **L'évaluation formative** : Elle a pour objectif de suivre l'évolution des compétences des apprenants au cours du processus d'apprentissage. Elle permet de repérer les difficultés rencontrées par l'apprenant, d'ajuster l'enseignement en conséquence et de fournir un retour immédiat sur ses performances. L'évaluation formative s'effectue de manière continue et offre aux apprenants la possibilité de s'améliorer au fur et à mesure du processus d'apprentissage. Selon Meunier (2019), « l'évaluation formative n'a pas pour vocation de juger l'apprenant, mais de l'accompagner dans ses progrès ».
- **L'évaluation sommative** : Contrairement à l'évaluation formative, l'évaluation sommative intervient généralement à la fin d'un cycle d'apprentissage, dans le but de mesurer les acquis d'un apprenant à un moment donné. Elle est souvent utilisée pour attribuer une note ou un certificat, comme les examens de fin de semestre ou les tests de certification (DELF, DALF). Cette forme d'évaluation permet de déterminer si l'apprenant a atteint les objectifs fixés au début du processus pédagogique. Pour Lemaire (2015), « l'évaluation sommative a une fonction de validation et de certification des connaissances et compétences acquises par l'apprenant ».

### **1.2. Les objectifs de l'évaluation dans l'enseignement des langues**

L'évaluation en didactique des langues sert plusieurs objectifs pédagogiques et didactiques, qui vont au-delà de la simple mesure des connaissances.

- **Mesure des compétences langagières** : L'évaluation permet de mesurer les compétences d'un apprenant dans différentes dimensions de la langue : la compréhension orale, la production orale, la compréhension écrite, et la production écrite. Ces quatre compétences langagières sont interconnectées et sont au cœur de

l'évaluation des apprenants en FLE. Comme le souligne Puren (2016), « l'évaluation doit prendre en compte l'intégralité des compétences langagières de l'apprenant, afin de donner une image complète de ses capacités communicatives ».

- **Amélioration continue de l'enseignement** : L'évaluation permet aussi à l'enseignant de se rendre compte de l'efficacité de sa pédagogie et d'adapter ses méthodes d'enseignement. Une évaluation bien conduite est donc aussi un outil de réflexion pour l'enseignant sur sa propre pratique. Selon Cassany (2012), « l'évaluation, loin de se limiter à une finalité de certification, doit permettre à l'enseignant de faire évoluer ses pratiques pédagogiques en fonction des retours qu'elle génère ».
- **Motivation et engagement des apprenants** : L'évaluation peut également jouer un rôle majeur dans la motivation des apprenants. Une évaluation formative bien menée peut renforcer la confiance en soi de l'apprenant et l'inciter à poursuivre ses efforts, tout en lui fournissant des repères sur son évolution. Selon Vaillancourt (2018), « une évaluation claire et transparente peut servir de moteur à l'engagement de l'apprenant, en l'aidant à voir ses progrès et en l'incitant à persévérer ».

### 1.3. Le rôle de l'évaluation dans le FLE

Dans le contexte spécifique du FLE, l'évaluation présente des enjeux particuliers. L'une des spécificités de l'évaluation en FLE réside dans le fait qu'elle ne se limite pas à évaluer la maîtrise des structures grammaticales ou lexiques, mais inclut également une dimension interculturelle et pragmatique. L'apprenant doit non seulement être capable de produire un discours correct sur le plan linguistique, mais aussi de l'adapter à différents contextes de communication.

- **Évaluation des compétences linguistiques et culturelles** : Le FLE ne se contente pas de l'évaluation des compétences linguistiques. L'apprenant est également évalué sur sa capacité à utiliser la langue de manière adaptée à des contextes sociaux et culturels spécifiques. « L'évaluation en FLE s'inscrit dans une approche communicative et interculturelle, où la compétence culturelle est tout aussi importante que la compétence linguistique » (Berthet, 2019).
- **Tests de certification et examens** : Le rôle de l'évaluation en FLE est également de certifier les compétences des apprenants à travers des tests standardisés, tels que le DELF (Diplôme d'Études en Langue Française) et le DALF (Diplôme Approfondi de

Langue Française). Ces tests évaluent l'ensemble des compétences langagières et permettent de mesurer le niveau de l'apprenant selon des critères internationaux.

- **L'évaluation de la production orale** : Une autre particularité de l'évaluation en FLE est l'importance de la production orale. L'oral, dans un contexte FLE, nécessite une évaluation qui va au-delà de la simple correction des erreurs de prononciation ou de grammaire ; elle inclut également la capacité à interagir de manière fluide, à comprendre et à répondre à des sollicitations orales dans un cadre social spécifique. Selon Goh (2017), « l'évaluation orale est l'un des défis majeurs dans l'enseignement du FLE, car elle implique de mesurer la capacité de l'apprenant à s'exprimer de manière spontanée et appropriée dans des situations variées ».

En conclusion, l'évaluation en didactique des langues, et spécifiquement dans le cadre du FLE, joue un rôle fondamental dans l'accompagnement de l'apprentissage, en permettant non seulement de mesurer les compétences acquises, mais aussi d'orienter et de réajuster l'enseignement pour mieux répondre aux besoins des apprenants. L'évaluation, bien qu'elle soit souvent perçue comme un moment de validation, doit avant tout être un outil d'accompagnement et d'amélioration continue pour l'apprenant comme pour l'enseignant.

## **2 L'intelligence artificielle : définition, types et applications en didactique des langues**

### **2.1 Définition de l'intelligence artificielle**

L'intelligence artificielle (IA) désigne un ensemble de méthodes, de théories et de techniques permettant à des machines ou programmes informatiques d'imiter certaines capacités de l'intelligence humaine. Ces capacités incluent la compréhension du langage, l'apprentissage, la reconnaissance de formes, le raisonnement ou encore la prise de décision.

Ainsi, Jean-Gabriel Ganascia (2017) définit l'IA comme « un ensemble de théories et de techniques mises en œuvre en vue de réaliser des machines capables de simuler l'intelligence humaine ». Autrement dit, l'IA tente de reproduire des processus cognitifs humains dans un but opérationnel : résoudre des problèmes, répondre à des questions, ou encore analyser des données à grande échelle.

Selon Cédric Villani (2018), l'IA représente « un outil au service de l'homme, capable de réaliser des analyses complexes sur de très grandes masses de données ». Elle permet, par exemple, de traiter rapidement des informations issues de multiples sources et de proposer des

solutions adaptées à une situation pédagogique donnée. Cela rend son usage particulièrement pertinent dans des contextes d'enseignement nécessitant une adaptation constante aux profils des apprenants.

L'intelligence artificielle fonctionne principalement grâce à des algorithmes. Un algorithme est une suite d'instructions précises que l'ordinateur exécute pour traiter des données. Lorsqu'une IA est entraînée, elle apprend à partir de nombreuses données (textes, images, sons, etc.) en identifiant des régularités ou des motifs. On parle alors d'apprentissage automatique (*machine learning*). Plus le système est entraîné sur des exemples variés et nombreux, plus il devient performant dans ses réponses.

Dans les systèmes dits d'apprentissage profond (*deep learning*), des réseaux de neurones artificiels sont utilisés. Inspirés du fonctionnement du cerveau humain, ces réseaux traitent les informations à travers plusieurs couches successives, ce qui leur permet d'effectuer des tâches complexes comme la reconnaissance d'images ou la compréhension de phrases longues.

En éducation, l'IA ne possède pas de conscience ou d'intelligence autonome ; elle fonctionne sur la base d'algorithmes créés par des humains, capables d'apprendre à partir de données, de s'ajuster à des situations nouvelles et de répondre à des requêtes variées. Elle peut ainsi assister les enseignants et soutenir les apprenants, sans pour autant se substituer à l'intervention humaine.

## **2.2 Les principaux types d'intelligence artificielle**

L'IA peut être classée selon deux critères principaux : le niveau de performance et l'approche technologique adoptée.

### **2.2.1. Selon le niveau de performance :**

- **IA faible (ou étroite)** : Elle est spécialisée dans une tâche bien définie (ex. : correction grammaticale, reconnaissance vocale). Elle ne peut pas s'adapter à des contextes différents sans intervention humaine. C'est la forme la plus répandue aujourd'hui.
- **IA forte (ou générale)** : Théorique pour le moment, elle représenterait une machine dotée de la capacité d'apprendre et de comprendre n'importe quelle tâche intellectuelle humaine. Elle fait l'objet de débats scientifiques, techniques et éthiques.



### 2.2.2. Selon l'approche technologique :

- **IA symbolique** : Elle repose sur des règles explicites formulées par des experts. Elle convient bien aux systèmes devant justifier leur raisonnement (ex. : diagnostic médical, aides à la décision).
- **IA connexionniste** : Inspirée du fonctionnement du cerveau, cette approche utilise des réseaux de neurones artificiels qui apprennent à partir d'exemples. Elle est à la base du *deep learning*.
- **IA hybride** : Elle combine règles explicites et apprentissage automatique, offrant des systèmes à la fois adaptables et interprétables.

### 2.3 Types d'IA utilisés en didactique des langues

Dans l'enseignement/apprentissage du français langue étrangère (FLE), plusieurs formes d'IA ont été progressivement intégrées pour personnaliser l'apprentissage, automatiser certaines tâches et offrir un accompagnement individualisé aux apprenants.

- **L'apprentissage automatique (machine learning)** : Cette technologie permet à des applications d'apprendre des comportements ou des erreurs des utilisateurs pour adapter en conséquence les contenus pédagogiques. Des outils comme Duolingo ou Lingvist ajustent ainsi les exercices selon le niveau et les performances de l'utilisateur.
- **Le traitement automatique du langage naturel (TALN)** : Il permet à une machine de traiter le langage humain de manière efficace. Grâce à lui, des outils comme ChatGPT, Grammalecte, ou Antidote peuvent corriger des textes, proposer des reformulations ou répondre à des questions linguistiques.
- **Les systèmes de tutorat intelligents** : Ces dispositifs analysent les besoins des apprenants et adaptent automatiquement les parcours proposés. Ils offrent un suivi personnalisé et des rétroactions ciblées, comme les plateformes Knewton ou ALEKS.
- **Les agents conversationnels** : Ils permettent de simuler des échanges oraux ou écrits dans la langue cible, facilitant ainsi l'apprentissage en contexte. Ce sont des assistants virtuels capables de dialoguer avec l'apprenant en lui posant des questions, en corrigeant ses réponses ou en lui proposant des jeux de rôle interactifs.

- **L'analyse prédictive et les *learning analytics*** : Ces outils analysent les traces d'apprentissage (réponses, temps passé, fréquence d'erreurs) pour détecter les difficultés, alerter l'enseignant et ajuster les contenus de manière anticipée.

Comme le rappellent Alberio et Poteaux (2020) dans *Numérique et apprentissage des langues*, ces technologies, si elles sont bien intégrées, permettent de « diversifier les pratiques pédagogiques et d'accompagner les apprentissages dans une perspective plus individualisée, tout en renforçant le rôle de l'enseignant comme médiateur et guide ».

## **2.3. L'introduction de l'intelligence artificielle dans l'éducation**

L'avènement de l'intelligence artificielle (IA) dans le domaine éducatif marque une transformation significative des pratiques pédagogiques. Elle se manifeste à travers une redéfinition des méthodes d'enseignement, d'apprentissage et surtout d'évaluation. Dans un contexte où l'enseignement des langues étrangères, notamment le français langue étrangère (FLE), cherche à répondre aux besoins de plus en plus individualisés des apprenants, l'IA se présente comme un levier puissant de personnalisation et de réactivité.

### **2.3.1. L'émergence de l'IA dans l'univers éducatif**

L'intégration de l'IA dans le champ éducatif n'est pas un phénomène récent, mais sa montée en puissance s'est accélérée avec le développement des technologies numériques. L'intelligence artificielle, selon Jean-Marie Mialaret (2018), « est capable de proposer une évaluation adaptative, c'est-à-dire une évaluation qui évolue en fonction des réponses de l'apprenant », soulignant ainsi sa capacité à ajuster les contenus et les modalités de l'évaluation en temps réel.

Cette évaluation adaptative permet de mieux répondre aux profils variés des apprenants, en s'appuyant sur l'analyse de données issues des interactions entre l'élève et le dispositif numérique. L'IA peut alors générer un parcours d'apprentissage personnalisé, en identifiant les zones de fragilité de l'apprenant et en lui proposant des exercices ciblés pour y remédier. Elle offre également une capacité de suivi longitudinal, permettant d'observer l'évolution des compétences sur le long terme.

### **2.3.2. Les apports concrets de l'IA à l'évaluation en langues**

Dans le cadre spécifique de l'enseignement/apprentissage des langues, l'IA présente plusieurs avantages opérationnels :

- **Analyse automatique des productions écrites** : des outils d'analyse sémantique et syntaxique permettent d'identifier les erreurs, de suggérer des corrections et de noter objectivement des copies, notamment sur des critères de grammaire, vocabulaire, cohérence textuelle, etc. Cela permet de soulager l'enseignant de certaines tâches répétitives tout en garantissant un retour rapide à l'apprenant.
- **Feedback immédiat et individualisé** : l'IA permet un retour quasi instantané sur les réponses de l'apprenant. Elle est en mesure de fournir des explications détaillées, de réorienter les exercices, et même de recommander des ressources complémentaires. Pour Bouvier et Trebbi (2020), « un feedback bien structuré est essentiel pour la progression de l'apprenant, et l'automatisation de ce feedback par l'IA peut en augmenter la fréquence et la précision ».
- **Simulation d'interactions orales** : des chatbots ou avatars dotés d'IA permettent aux apprenants de s'entraîner à l'oral en simulant des conversations authentiques. Bien que ces dispositifs aient encore des limites en matière de compréhension contextuelle, ils offrent un terrain d'entraînement motivant et personnalisé.
- **Outils d'autoévaluation** : l'IA favorise aussi le développement de la capacité d'autoévaluation chez les apprenants. En les incitant à comparer leurs productions à des modèles de référence et en leur fournissant des critères clairs, elle développe leur autonomie et leur esprit critique, comme le souligne Legros (2021), pour qui « l'évaluation ne peut être un outil de formation qu'à condition d'être comprise et appropriée par l'apprenant ».

### 2.3.3. Les limites et les enjeux de l'IA dans l'éducation

Malgré ses apports notables, l'IA soulève également des préoccupations. Sur le plan pédagogique, il est important de rappeler que l'apprentissage des langues est un processus profondément humain, dans lequel l'interaction sociale, l'émotion et la médiation culturelle jouent un rôle central. Or, comme le rappelle Lemoine (2019), « l'usage croissant des technologies dans l'enseignement ne doit pas conduire à une déshumanisation des pratiques éducatives ».

Par ailleurs, il est essentiel de distinguer intelligence humaine et intelligence artificielle, non seulement sur le plan technique, mais aussi dans leurs implications pédagogiques. L'intelligence humaine est fondée sur la conscience, l'émotion, la subjectivité, la créativité et la capacité à s'adapter à des situations inédites. Elle évolue dans un contexte

social et culturel riche, mobilisant l'intuition, l'empathie et le jugement. À l'inverse, l'intelligence artificielle repose sur des calculs statistiques et des algorithmes qui traitent des données à grande échelle. Elle fonctionne de manière logique, prédictive, mais sans conscience ni sensibilité.

Comme le rappellent Bachimont et Mille (2019), « l'IA peut simuler certains comportements intelligents, mais elle ne comprend pas ; elle ne fait que manipuler des signes à partir de corrélations statistiques ». Cette distinction est fondamentale, car elle montre que l'IA n'est pas capable de faire sens de manière autonome, et que l'enseignant, en tant qu'acteur humain, demeure indispensable pour interpréter, contextualiser et accompagner les apprentissages.

D'un point de vue éthique, plusieurs questions se posent : que devient la confidentialité des données des apprenants ? Qui contrôle les algorithmes d'évaluation ? Quels biais peuvent exister dans les systèmes de notation automatique ? De plus, les algorithmes sont souvent construits à partir de corpus de données partiels, biaisés ou culturellement situés, ce qui peut générer des résultats discriminants, notamment pour les apprenants non natifs ou issus de cultures différentes.

À cela s'ajoute la question de la transparence des systèmes : la plupart des outils d'IA sont des « boîtes noires », dont le fonctionnement interne n'est pas accessible aux enseignants ni aux apprenants. Cela pose un défi majeur à la pédagogie critique et à la responsabilisation des acteurs éducatifs.

Enfin, les outils d'IA, bien qu'efficaces pour des tâches précises, ne sont pas infaillibles. Leur capacité à comprendre des productions nuancées, à intégrer le contexte culturel ou à juger de la pertinence d'un discours reste limitée. Ils peinent à reconnaître l'implicite, les jeux de langage ou les subtilités stylistiques propres à chaque apprenant. Ainsi, l'évaluation automatique ne peut remplacer l'analyse fine d'un enseignant expérimenté. Elle doit plutôt être pensée comme un complément à l'action humaine, un outil d'assistance et non de substitution. Le risque d'une standardisation excessive des apprentissages et des évaluations est réel si l'IA est utilisée sans discernement pédagogique.

#### **2.3.4. Vers une complémentarité entre IA et pédagogie humaine**

L'intégration de l'IA dans l'évaluation des compétences linguistiques en FLE appelle à un repositionnement du rôle de l'enseignant. Loin de le marginaliser, l'IA peut lui permettre

de recentrer son action sur l'accompagnement, la différenciation pédagogique et la remédiation ciblée.

Pour Bessonnat (2022), « l'enseignant reste l'acteur principal de la médiation pédagogique. L'intelligence artificielle doit renforcer sa capacité à intervenir avec pertinence et bienveillance auprès de ses apprenants ».

En somme, l'IA offre des possibilités intéressantes pour l'évaluation en FLE, notamment par l'automatisation des tâches répétitives et la personnalisation de l'apprentissage. Mais son efficacité dépendra toujours d'une conception pédagogique rigoureuse, d'un encadrement humain avisé, et d'une réflexion continue sur les finalités éducatives.

### **3. Les principes de l'évaluation dans l'enseignement des langues**

L'évaluation constitue un pilier central de toute démarche pédagogique, particulièrement dans l'enseignement/apprentissage des langues. Loin d'être une simple procédure de mesure des acquis, elle s'inscrit dans une logique de régulation, de motivation et d'accompagnement des parcours d'apprentissage. Dans cette section, nous aborderons les objectifs fondamentaux de l'évaluation dans le domaine des langues, tout en soulignant les principes directeurs qui garantissent sa qualité, son efficacité et son équité.

#### **3.1. Objectifs de l'évaluation dans l'enseignement des langues**

L'évaluation en langues a pour vocation de vérifier les compétences langagières développées par l'apprenant, mais elle doit également servir à orienter son apprentissage et à nourrir la pratique de l'enseignant. Elle remplit ainsi plusieurs fonctions :

- **Fonction diagnostique** : en début de parcours, l'évaluation permet d'identifier les acquis préalables de l'apprenant, de repérer ses besoins, et d'ajuster les contenus et les modalités d'enseignement en conséquence.
- **Fonction formative** : elle accompagne le processus d'apprentissage, en fournissant un feedback régulier et structuré. Cette fonction est cruciale dans l'enseignement des langues, où la progression repose souvent sur une prise de conscience des erreurs et une auto-régulation des pratiques langagières.

- **Fonction sommative** : elle intervient à la fin d'une séquence d'enseignement pour mesurer les acquis de manière plus formelle, souvent à des fins de certification ou de validation.

Comme le souligne Beacco (2007), « l'évaluation doit être intégrée à l'enseignement des langues comme un instrument de pilotage des apprentissages et non comme une sanction ». Elle sert donc autant à améliorer l'enseignement qu'à faire progresser l'élève.

### **3.2. Les principes fondamentaux d'une évaluation de qualité**

Pour qu'une évaluation soit juste, efficace et utile, elle doit respecter certains principes fondamentaux qui font consensus dans la recherche en didactique des langues.

#### **3.2.1. La validité**

La validité d'une évaluation renvoie à sa capacité à mesurer effectivement ce qu'elle prétend mesurer. Autrement dit, une épreuve doit être en adéquation avec les objectifs d'apprentissage visés. Dans le contexte du FLE, il s'agit par exemple de s'assurer qu'une évaluation orale mesure bien la capacité à interagir ou à s'exprimer, et non la simple capacité de mémorisation de structures.

Selon Alderson, Clapham et Wall (1995), « une évaluation valide est celle qui rend compte de manière fidèle des compétences langagières définies dans le curriculum ou le référentiel ». L'alignement entre les objectifs, les contenus et les modalités d'évaluation est donc crucial.

#### **3.2.2. La fiabilité**

La fiabilité désigne la constance des résultats d'une évaluation : si plusieurs évaluateurs corrigent une même production, ou si un même élève est évalué à plusieurs reprises, les résultats doivent rester comparables. Cela suppose des critères d'évaluation clairs, partagés, et appliqués de manière cohérente.

Pour M. Byram (2000), « la fiabilité repose sur la transparence des critères et sur une formation solide des évaluateurs, condition indispensable pour garantir l'objectivité du jugement ». La mise en place de grilles critériées et de barèmes précis favorise cette fiabilité.

#### **3.2.3. L'équité**

Une évaluation équitable garantit l'égalité des chances pour tous les apprenants, quels que soient leur origine culturelle, leur parcours scolaire ou leur genre. Elle évite les biais culturels, linguistiques ou sociaux, et prend en compte les besoins spécifiques de certains apprenants.

Meyer (2020) insiste sur cette exigence : « Une évaluation juste est celle qui respecte les principes de validité, de fiabilité et d'équité, tout en ayant pour but de faire progresser l'apprenant ». Cela implique aussi d'adapter les supports et les consignes, et de proposer des modalités variées (écrites, orales, numériques) pour permettre à chacun d'exprimer ses compétences.

### **3.2.4. L'impact pédagogique**

L'évaluation ne doit pas être neutre : elle influence la manière dont les élèves apprennent. C'est ce qu'on appelle le *washback effect*, ou effet rétroactif de l'évaluation. Lorsqu'elle est bien conçue, l'évaluation peut encourager la réflexion métacognitive, la prise d'initiative et la motivation de l'apprenant. Elle devient ainsi un moteur d'apprentissage.

Black et Wiliam (1998), dans leur rapport fondateur, soulignent que « l'évaluation formative, lorsqu'elle est pratiquée régulièrement, a un impact significatif sur l'apprentissage, notamment chez les élèves en difficulté ». L'idée est donc d'utiliser l'évaluation non pour classer, mais pour accompagner et faire progresser.

### **3.3. Vers une évaluation formative, inclusive et réflexive**

À la lumière de ces principes, les pratiques évaluatives en langues évoluent aujourd'hui vers des approches plus formatives, inclusives et réflexives. L'évaluation formative se distingue par son objectif d'accompagner l'apprenant dans sa progression continue plutôt que de simplement certifier un niveau de compétence à un moment donné. Comme le précisent Black et Wiliam (2009), « l'évaluation formative est l'un des moyens les plus efficaces pour améliorer les résultats des élèves », car elle permet d'identifier les besoins spécifiques des apprenants et d'ajuster les démarches pédagogiques en conséquence.

Dans cette dynamique, l'enseignant devient un médiateur de l'évaluation : il guide, conseille, et crée des espaces de dialogue où l'apprenant peut comprendre ses erreurs et progresser. L'apprenant, quant à lui, n'est plus simple destinataire de la note, mais acteur engagé dans son processus d'évaluation. Cette approche répond aux principes du *learner-centered assessment* tels que décrits par Stiggins (2005), pour qui « l'évaluation doit être perçue par l'apprenant comme une opportunité de croissance personnelle et professionnelle ».

L'inclusivité est un autre pilier fondamental de cette évolution. Une évaluation inclusive vise à reconnaître la diversité des profils d'apprenants en prenant en compte leurs besoins spécifiques, leurs rythmes d'apprentissage et leurs contextes socioculturels. Comme l'affirme

Brookhart (2013), « une évaluation équitable et inclusive doit offrir à chaque élève des chances égales de démontrer ce qu'il sait et sait faire, sans être entravé par des facteurs extérieurs à l'objet évalué ». Cela implique par exemple l'adaptation des consignes, l'utilisation de supports variés, et le recours à des méthodes d'évaluation alternatives (portfolios, projets, autoévaluations).

L'intégration d'outils numériques et de l'intelligence artificielle dans les pratiques évaluatives ouvre de nouvelles perspectives pour favoriser cette évaluation formative et inclusive. Les plateformes adaptatives permettent de personnaliser les parcours d'évaluation, en fonction du niveau et du profil de l'apprenant. Cependant, comme le rappellent Popenici et Kerr (2017), « l'usage de l'intelligence artificielle en éducation doit être réfléchi et encadré afin de préserver l'autonomie de l'apprenant et de respecter les principes éthiques fondamentaux ». L'évaluation assistée par IA ne doit jamais devenir un processus automatisé déshumanisé : elle doit rester au service d'un projet pédagogique centré sur l'humain.

Enfin, l'évaluation réflexive prend une place croissante dans les dispositifs d'enseignement des langues. Elle invite l'apprenant à porter un regard critique sur son propre apprentissage, à identifier ses stratégies efficaces, et à élaborer des plans d'action pour progresser. Selon Schön (1983), « la pratique réflexive est essentielle pour tout apprentissage en profondeur », car elle engage l'apprenant dans une démarche métacognitive durable.

En somme, une évaluation de qualité dans l'enseignement des langues repose aujourd'hui sur quatre piliers essentiels : la validité, la fiabilité, l'équité et l'impact pédagogique. Elle ne se réduit pas à une mesure statique des performances, mais s'inscrit dans une perspective dynamique d'accompagnement, de valorisation des parcours d'apprentissage et d'émancipation des apprenants.

## **4. Innovations : l'apport de l'intelligence artificielle à l'évaluation en FLE**

### **4.1. Les apports de l'IA à l'évaluation linguistique**

Dans l'enseignement/apprentissage du français langue étrangère (FLE), l'intelligence artificielle (IA) ouvre de nouvelles perspectives pour améliorer les pratiques évaluatives. Trois principales contributions peuvent être identifiées :

#### **4.1.1. Correction automatisée des erreurs grammaticales ou lexicales**

Les outils fondés sur l'IA permettent de détecter et de corriger automatiquement des erreurs dans les productions écrites des apprenants. Des logiciels tels que Grammalecte ou



LanguageTool exploite des moteurs de traitement du langage naturel (NLP) pour proposer des suggestions grammaticales, lexicales ou orthographiques adaptées au français. Selon Burstein et Chodorow (2010), les systèmes d'analyse linguistique automatiques peuvent « fournir une rétroaction immédiate sur la grammaire et l'usage linguistique, contribuant à un apprentissage plus autonome et itératif » (*Burstein & Chodorow, 2010*).

#### **4.1.2. Analyse sémantique pour évaluer la cohérence textuelle**

L'IA peut aussi aller au-delà de la simple correction formelle pour analyser la cohérence logique et l'organisation discursive d'un texte. Des systèmes d'analyse sémantique latente (LSA) ou d'embedding textuel (type Word2Vec ou BERT) sont capables de repérer des ruptures dans le fil du discours ou des incohérences argumentatives (Landauer, Foltz & Laham, 1998). Ces outils permettent ainsi une évaluation plus fine de la production écrite, prenant en compte la dimension pragmatique du langage, essentielle en FLE.

#### **4.1.3. Reconnaissance vocale pour évaluer l'oralité**

La reconnaissance automatique de la parole (ASR) est de plus en plus exploitée pour évaluer la production orale. Elle permet d'analyser :

- La **prononciation**,
- L'**intonation**,
- La **fluidité** du discours.

Des applications comme Speak & Improve (développée par l'Université de Cambridge) ou des modules de plateformes comme Duolingo utilisent des algorithmes de reconnaissance vocale entraînés spécifiquement pour évaluer des locuteurs non natifs. Comme le souligne Derwing et Munro (2005), « la perception de la prononciation est un élément clé de l'évaluation de la compétence orale, et l'automatisation peut aider à objectiver ce processus » (*Derwing & Munro, 2005*).

#### **Exemples concrets :**

- **Grammalecte** : extension libre spécialisée pour la langue française, intégrant une correction grammaticale assistée par IA.
- **LanguageTool** : correcteur multilingue, doté de règles de grammaire complexes et de détections sémantiques en français.

- **Speak & Improve** : plateforme utilisant l'IA pour évaluer la compétence orale des apprenants de FLE selon des critères CECRL.

## 4.2. Vers une évaluation plus personnalisée et adaptative

L'un des apports majeurs de l'IA à l'évaluation en FLE réside dans la capacité de proposer des processus **plus flexibles** et **centrés sur l'apprenant**.

### 4.2.1. Tests adaptatifs basés sur les réponses de l'apprenant

Grâce aux algorithmes de machine learning, il est possible de concevoir des **tests adaptatifs** : la difficulté des questions s'ajuste dynamiquement selon les réponses de l'apprenant, offrant ainsi une mesure plus précise de son niveau réel. Par exemple, le Duolingo English Test utilise ce principe pour ses évaluations certifiantes. Comme l'expliquent Bernstein, Van Moere et Cheng (2010), « les tests adaptatifs assistés par IA augmentent la validité de l'évaluation tout en réduisant la durée du test » (*Bernstein, Van Moere, & Cheng, 2010*).

### 4.2.2. Feedbacks instantanés et individualisés

L'IA permet de générer des **commentaires personnalisés** en temps réel, portant non seulement sur les erreurs, mais aussi sur les stratégies d'amélioration possibles. Les recherches de Heift et Schulze (2007) montrent que des rétroactions spécifiques et immédiates « *favorisent un apprentissage linguistique plus efficace que des corrections différées ou génériques* » (Heift & Schulze, 2007).

### 4.2.3. Analyse des profils d'apprentissage

Enfin, des systèmes basés sur l'analyse de données (learning analytics) peuvent tracer les performances des apprenants sur la durée, identifier leurs points forts et leurs lacunes, et proposer des parcours d'apprentissage sur mesure. Selon Ifenthaler et Yau (2020), l'utilisation de l'IA pour personnaliser l'évaluation « contribue à développer des parcours d'apprentissage plus engagés et plus autonomes » (*Ifenthaler & Yau, 2020*).

## 4.3 Limites de l'évaluation par l'IA en FLE

Cependant, plusieurs précautions doivent être prises :

- Biais d'entraînement des IA sur des corpus non représentatifs de la diversité linguistique réelle des apprenants.

- Réduction de la complexité humaine dans l'interprétation de la communication interculturelle.
- Opacité de certains algorithmes utilisés pour noter ou corriger.

### **3.4 Vers une complémentarité entre l'enseignant et l'intelligence artificielle**

L'IA ne doit pas remplacer l'enseignant-évaluateur, mais plutôt :

- Alléger les tâches répétitives (correction simple, tri automatique des erreurs).
- Soutenir la prise de décision pédagogique.
- Offrir des outils d'accompagnement pour affiner l'analyse qualitative faite par l'enseignant.

En somme, l'avenir de l'évaluation en FLE repose sur une synergie intelligente entre technologie et expertise humaine.

## **Conclusion**

L'examen des cadres théoriques liés à la compréhension de l'oral en FLE, à l'évaluation formative, ainsi qu'aux apports récents de l'intelligence artificielle dans l'enseignement-apprentissage des langues, permet de dégager un ensemble de constats fondamentaux pour aborder sereinement l'expérimentation menée dans ce mémoire.

Tout d'abord, la compréhension de l'oral apparaît comme une compétence linguistique hautement stratégique, mais aussi particulièrement exigeante pour les apprenants non natifs. Elle suppose une maîtrise de plusieurs dimensions : linguistique (vocabulaire, morphosyntaxe), phonétique (reconnaissance des sons, intonation), cognitive (traitement en temps réel), et sociolinguistique (prise en compte des contextes, des implicites, des variations culturelles). Ces paramètres rendent son évaluation d'autant plus délicate, car elle ne peut se limiter à des réponses correctes ou incorrectes : elle doit viser à cerner les processus de compréhension, les obstacles rencontrés, les interprétations possibles et la manière dont l'apprenant mobilise ses connaissances pour construire du sens.

Ensuite, les pratiques évaluatives en classe de FLE doivent aujourd'hui évoluer vers des formes plus formatives, interactives et engageantes. L'évaluation ne saurait se réduire à un jugement terminal ; elle doit participer pleinement à la dynamique d'apprentissage, en accompagnant l'étudiant dans la construction de ses compétences. Dans ce cadre, le retour

donné à l'apprenant joue un rôle central : il doit être rapide, clair, personnalisé et porteur de sens.

C'est précisément sur ce terrain que l'intelligence artificielle – et notamment les modèles de traitement du langage comme ChatGPT – suscite un intérêt croissant. En permettant une analyse instantanée de la production linguistique et un retour automatisé mais structuré, ces outils représentent un levier prometteur, tant pour l'enseignant (qui gagne en temps et en réactivité) que pour l'étudiant (qui reçoit un feedback immédiat, souvent plus neutre et dédramatisé). Toutefois, comme toute innovation, leur intégration doit être pensée avec précaution : il ne s'agit pas de remplacer le rôle de l'enseignant, mais de compléter son action, dans un cadre méthodologique et éthique clairement défini.

À la lumière de ces éléments, l'expérimentation présentée dans la seconde partie de ce mémoire a pour ambition d'évaluer, sur le terrain, l'impact concret de l'intelligence artificielle dans une situation d'évaluation de la compréhension de l'oral. Elle a été menée auprès d'étudiants de troisième année licence au département de français de l'Université de Biskra, répartis en deux groupes distincts : l'un évalué selon une méthode traditionnelle, l'autre bénéficiant d'une correction assistée par IA.

Cette démarche vise à confronter les apports théoriques à une réalité pédagogique, en analysant les effets de l'IA sur la qualité du feedback, l'engagement des apprenants, et l'efficacité globale du dispositif. Elle permettra également de mettre en évidence les perceptions des étudiants, les difficultés éventuelles, et les conditions nécessaires à une intégration réussie de ces technologies dans l'enseignement supérieur algérien.

La partie suivante présentera en détail le cadre de l'expérimentation, la méthodologie adoptée, les résultats observés et les pistes de réflexion qu'ils ouvrent pour l'avenir de la didactique du FLE à l'ère du numérique.

**CHAPITRE II : ÉTUDE EXPERIMENTALE SUR  
L'EVALUATION DE LA COMPREHENSION ORALE  
ASSISTEE OU NON PAR IA**

## Introduction

À l'ère du numérique, les technologies d'intelligence artificielle investissent progressivement le champ de l'enseignement/apprentissage des langues, en particulier dans les domaines de l'évaluation et de la rétroaction pédagogique. Si les apports théoriques sur l'usage de l'IA dans le FLE ont été largement débattus dans la première partie de ce mémoire, il importe désormais de confronter ces réflexions à la réalité du terrain, afin d'en mesurer la pertinence, les limites et les implications concrètes.

Ce second chapitre propose une étude expérimentale menée auprès d'un public ciblé : les étudiants de troisième année licence du département de français de l'Université de Biskra, engagés dans un cours de compréhension de l'oral. L'objectif est de comparer deux modalités d'évaluation de la compréhension orale : l'une traditionnelle (feedback manuel fourni par l'enseignant), l'autre innovante (feedback généré avec l'assistance de l'intelligence artificielle, via un outil comme ChatGPT).

L'expérimentation vise à répondre à une série de questions pédagogiques essentielles :

- L'intelligence artificielle peut-elle améliorer la qualité et la rapidité du retour donné à l'apprenant ?
- Son usage influence-t-il la motivation, la perception de l'évaluation et la progression des étudiants ?
- Quelles sont les conditions nécessaires à une intégration pertinente et éthique de l'IA dans le contexte universitaire algérien ?

Ce chapitre s'organise en plusieurs sections. Il commence par une présentation du contexte de l'étude (cadre institutionnel, publics concernés), puis détaille la méthodologie adoptée (dispositif expérimental, outils, corpus). Les résultats seront ensuite analysés à la lumière des variables identifiées, avant de dégager les enseignements principaux de cette expérimentation, en tenant compte à la fois des apports observés et des limites rencontrées.

Cette étude, bien que circonscrite, ambitionne de contribuer à la réflexion actuelle sur les usages pertinents de l'intelligence artificielle en didactique du FLE, en articulant innovation technologique et rigueur méthodologique dans une perspective locale et contextualisée.

## **1. Cadre méthodologique**

Cette étude expérimentale vise à évaluer l'apport de l'intelligence artificielle, notamment ChatGPT, dans l'analyse et la correction des réponses à un questionnaire de compréhension de l'oral en FLE. Afin de comparer l'efficacité d'une évaluation classique à celle assistée par IA, deux groupes d'étudiants seront constitués. Le protocole retenu s'organise autour de plusieurs phases pédagogiques communes aux deux groupes, avec une différenciation dans le mode de correction et de restitution des résultats.

### **1.1 Objectifs de l'expérimentation**

- Présenter les objectifs principaux de l'expérimentation : comparer deux modalités d'évaluation (classique vs. automatisée par IA).
- Souligner l'intérêt didactique et pédagogique de l'étude dans un contexte FLE universitaire algérien.

### **1.2. Participants**

L'échantillon envisagé sera composé de deux groupes d'étudiants inscrits en troisième année licence de français (L3) à l'Université de Biskra. Le premier groupe constituera le groupe témoin, et le second, le groupe expérimental. Les deux groupes seront comparables en termes de niveau linguistique, d'âge, de formation antérieure et d'exposition au français, assurant une homogénéité relative dans le profil des participants.

### **1.2. Matériel utilisé**

Le dispositif expérimental reposera sur les éléments suivants :

- Une capsule vidéo extraite de TV5MONDE intitulée « Infographie – Les langues du monde », sélectionnée pour sa richesse thématique et sa brièveté adaptée au format pédagogique ;
- Un questionnaire Google Forms comprenant différents types de questions (QCM, repérage, question ouverte) destiné à évaluer la compréhension de la vidéo ;
- Des outils d'analyse différenciés selon les groupes : une correction manuelle classique pour le groupe témoin, et une évaluation assistée par IA via ChatGPT exploitée à partir d'un tableau Google Sheets pour le groupe expérimental.

### 1.3. Déroulement général

Les deux groupes suivront une séquence pédagogique identique, structurée en trois temps :

- Pré-visionnage : activités d'anticipation orale autour de la thématique ;
- Visionnage de la capsule vidéo (deux passages prévus) ;
- Post-visionnage : passation du questionnaire de compréhension.

La différence principale portera sur le mode d'évaluation : manuel pour le groupe témoin, et assisté par l'intelligence artificielle pour le groupe expérimental.

### 1.4. Outil complémentaire d'observation : la grille

Afin de compléter l'analyse des résultats obtenus via le questionnaire et les modes de correction, une grille d'observation a été conçue et utilisée de manière identique pour les deux groupes (témoin et expérimental).

Cette grille vise à observer, pendant le déroulement de la séance d'évaluation, certains aspects comportementaux et organisationnels qui ne sont pas directement mesurables par les outils d'évaluation classiques. Elle permet ainsi de documenter la dynamique de classe, les réactions des étudiants face aux consignes, aux outils numériques, et au moment du retour évaluatif.

La grille a été élaborée sur la base des besoins suivants :

- Observer l'implication des étudiants pendant les trois phases de la séance (prévisionnage, visionnage, post-visionnage) ;
- Repérer les difficultés techniques (accès à l'outil, gestion du temps, autonomie) ;
- Évaluer la réaction des étudiants face au feedback différé (groupe témoin) ou automatisé et individualisé (groupe expérimental) ;
- Comparer objectivement les deux modalités d'évaluation au-delà des simples performances chiffrées.

Élément observé	Indicateurs	Observations
Implication des étudiants	Attention, prise de notes, questions posées	Exemple : majorité concentrée
Réaction face à la consigne	Compréhension, hésitation, reformulation demandée	Ex. : besoin de reformuler



<b>Élément observé</b>	<b>Indicateurs</b>	<b>Observations</b>
Autonomie dans la tâche	Capacité à répondre sans sollicitation constante	Autonomie moyenne
Comportement face à l'outil (smartphone/IA)	Difficulté technique, fluidité d'usage	Fluidité satisfaisante
Temps de réponse aux questions	Rapidité ou lenteur d'exécution	Moyenne : 7 à 10 minutes
Attitude face à la correction ou au feedback	Intérêt, indifférence, questionnement	Réaction faible (groupe témoin) ou engagement (groupe expérimental)
Interactions entre pairs	Discussions, entraide, partage de réponses	Présentes chez 5-6 étudiants
Problèmes techniques rencontrés	Connexion, batterie, outils	Aucun problème majeur signalé

Elle comporte une série d'indicateurs relatifs à :

- la participation active,
- la gestion de l'activité sur smartphone,
- l'interaction entre pairs,
- le respect du temps,
- les réactions face au feedback.

Les observations ont été prises en temps réel par l'enseignante-chercheuse, et résumées en fin de séance. Bien que qualitative, cette grille offre un appui concret pour interpréter les écarts observés entre les deux groupes au moment de la synthèse comparative.

## **2. Description des modalités expérimentales**

### **2.1. Protocole pour le groupe témoin : évaluation manuelle classique**

#### **2.1.1. Conditions matérielles**

L'évaluation du groupe témoin se déroulera dans une salle de cours classique, sans accès à une salle multimédia. Les étudiants utiliseront leurs smartphones personnels, éventuellement couplés à un haut-parleur central pour une écoute collective de la vidéo.

#### **2.1.2. Déroulement de la séance**

La séance se déroulera de la manière suivante :

- Pré-écoute : brève activité d'anticipation orale autour du thème des langues ;
- Visionnage de la capsule (deux fois), sur smartphone avec audio amplifié ;
- Remplissage du questionnaire via Google Forms immédiatement après le visionnage.

#### **2.1.3. Contenu du questionnaire**

Le questionnaire utilisé dans cette expérimentation a été conçu pour évaluer la compréhension orale des étudiants à partir de la vidéo « Infographie : les langues du monde » produite par TV5MONDE. Il comporte trois parties distinctes, correspondant aux étapes classiques d'une séquence de compréhension audiovisuelle : prévisionnage, visionnage et post-visionnage. Chaque partie cible des compétences spécifiques.

- Partie A – Prévisionnage (2 points)

Elle contient deux questions à choix multiples (QCM) permettant d'activer les connaissances préalables des apprenants sur la diversité linguistique dans le monde. Elle prépare les étudiants à la réception de la vidéo.

- Partie B – Visionnage (6 points)

Cette partie comprend cinq questions destinées à évaluer la compréhension précise du contenu vidéo, allant de l'extraction de faits chiffrés à l'identification d'informations spécifiques. Elle mêle des questions fermées (QCM ou réponses courtes) et une question de repérage de familles linguistiques.

- Partie C – Post-visionnage (2 points)

Elle propose une question ouverte visant à recueillir une réaction personnelle argumentée sur l'importance de la préservation des langues menacées. Elle évalue à la fois la pertinence du propos, la qualité de l'expression et l'esprit critique de l'apprenant.

Le barème global de l'évaluation est de 10 points. Cette structuration progressive permet d'observer les capacités des étudiants à mobiliser leurs connaissances, à extraire des informations pertinentes à l'écoute et à formuler une opinion personnelle en lien avec le thème abordé.

#### **2.1.4. Correction et évaluation**

Les réponses seront lues et corrigées manuellement par l'enseignante à partir de Google Forms. Le temps estimé pour la correction est d'environ 1h30 pour 25 étudiants. Les étudiants recevront une note sur 10, sans feedback écrit. Un retour oral différé sera prévu lors de la séance suivante.

### **2.2. Protocole pour le groupe expérimental : évaluation assistée par IA**

#### **2.2.1. Conditions identiques au groupe témoin**

Le groupe expérimental sera soumis aux mêmes conditions matérielles et pédagogiques que le groupe témoin : même vidéo, même questionnaire, même séquence didactique et même environnement.

#### **2.2.2. Procédure d'évaluation automatisée**

Les réponses collectées via Google Forms seront exportées sur Google Sheets, puis soumises à une analyse automatisée à l'aide de ChatGPT, selon une grille de critères prédéfinis :

- Pertinence de la réponse ;
- Cohérence syntaxique ;
- Clarté de la formulation.

Une vérification manuelle aléatoire sur environ 20 % des copies sera prévue pour contrôler la fiabilité de l'évaluation automatisée. Le temps global de traitement est estimé à moins de 15 minutes.

#### **2.2.3. Retour individualisé**

Chaque étudiant recevra un feedback généré automatiquement par l'IA comprenant :

- Un commentaire général sur sa performance ;
- Des remarques ciblées sur la syntaxe ou la pertinence ;

- Une reformulation optimisée de la réponse ouverte, si nécessaire.

Le feedback sera transmis dans un délai très court, permettant une prise de conscience immédiate des erreurs.

## **2.3. Comparaison des deux modalités**

### **2.3.1. Critères comparatifs**

Plusieurs critères seront retenus pour analyser l'écart entre les deux modalités :

- Temps de traitement et de correction ;
- Nature et qualité du feedback (présence ou absence de commentaires, personnalisation) ;
- Réception par les étudiants : motivation, engagement, satisfaction ;
- Faisabilité technique dans le contexte universitaire algérien (connectivité, équipement, accessibilité aux outils IA).

### **2.3.2. Synthèse des avantages et inconvénients**

Une analyse comparative sera proposée sous forme de tableau synthétique, accompagnée d'un commentaire interprétatif permettant de cerner les apports et les limites de l'évaluation assistée par IA dans un contexte d'enseignement/apprentissage du FLE.

## **3. Déroulement de l'expérimentation : Groupe témoin (évaluation classique)**

L'expérimentation menée avec le groupe témoin s'est déroulée dans des conditions pédagogiques simples, sans recours aux outils numériques avancés ni à une salle multimédia. L'ensemble de la séance a été conduit en classe ordinaire, avec l'utilisation exclusive de matériel papier et des smartphones personnels des étudiants.

### **3.1. Phase de pré-visionnage**

La séance a débuté par une brève activité d'anticipation orale, au cours de laquelle l'enseignante a invité les étudiants à formuler des hypothèses sur le contenu de la capsule vidéo à partir de son titre : « *Infographie : les langues du monde* », proposé par la chaîne francophone TV5MONDE. Quelques mots-clés génériques ont été inscrits au tableau (« langues », « diversité linguistique », « chiffres »), afin de stimuler la réflexion des apprenants sur le thème de la diversité linguistique mondiale.

### **3.2. Visionnage de la capsule vidéo**

Les étudiants ont été invités à visionner la vidéo deux fois consécutives sur leurs smartphones personnels, connectés à Internet. Le visionnage a été collectif, chacun sur son appareil, en utilisant les haut-parleurs ou les écouteurs disponibles. Aucune infrastructure multimédia n'était disponible dans la salle (pas d'ordinateur, ni projecteur, ni enceintes). La durée totale de la vidéo était d'environ 2 minutes 30. Entre les deux visionnages, un court temps d'échange a permis aux étudiants de signaler ce qu'ils avaient retenu ou les éléments perçus comme importants.

### **3.3. Passation du questionnaire de compréhension**

À l'issue du deuxième visionnage, un questionnaire imprimé sur support papier a été distribué aux étudiants. Ce questionnaire, identique à celui utilisé pour le groupe expérimental, comportait quatre questions, conçues pour évaluer plusieurs dimensions de la compréhension orale.

Les étudiants ont eu environ 10 minutes pour remplir le questionnaire. Une fois les copies récupérées, l'enseignante a clos la séance.

### **3.4. Traitement des réponses et évaluation**

La correction des copies a été réalisée manuellement par l'enseignante, en dehors du temps de classe. Chaque questionnaire a été traité individuellement, en consultant les réponses inscrites sur le papier. Le temps estimé pour la correction complète des 25 copies a été d'environ 1 heure 30 minutes, notamment en raison de la question ouverte, qui nécessitait une lecture attentive, bien que succincte, de chaque production écrite.

En raison de contraintes temporelles, aucun feedback détaillé n'a été rédigé pour les apprenants. Lors de la séance suivante, l'enseignante s'est contentée d'indiquer à chaque étudiant si ses réponses étaient correctes ou non, et de suggérer une correction linguistique rapide pour la formulation de la réponse ouverte. Les apprenants ont reçu une note globale sur 10, sans grille de critères explicite ni commentaire écrit individualisé.

## **4. Protocole pour le groupe expérimental : évaluation assistée par intelligence artificielle**

À la suite de l'expérimentation menée avec le groupe témoin, une deuxième expérimentation a été réalisée avec un groupe expérimental de même niveau (L3), composé également de 25 étudiants issus du département de français à l'université de Biskra. L'objectif était d'évaluer l'apport de l'intelligence artificielle dans le processus d'évaluation de la compréhension de l'oral en FLE, en utilisant un protocole identique, à l'exception de la phase de correction et de retour aux apprenants.

### **4.1. Conditions matérielles et pédagogiques**

L'expérimentation s'est déroulée dans des conditions strictement comparables à celles du groupe témoin. Aucun laboratoire multimédia n'étant disponible, les étudiants ont utilisé leurs propres smartphones pour accéder au questionnaire Google Forms. L'écoute de la capsule vidéo s'est faite via un haut-parleur externe connecté au smartphone de l'enseignante, dans une salle de cours classique.

La capsule vidéo utilisée est la même que pour le groupe témoin : TV5MONDE - Infographie : Les langues du monde

### **4.2. Déroulement de la séance**

La séance s'est articulée en trois phases distinctes, conformes à une approche pédagogique communicative :

- **Pré-écoute (5 min)** : L'enseignante a engagé une activité orale d'anticipation autour du thème des langues et du plurilinguisme, en demandant aux étudiants de citer les langues les plus parlées dans le monde, et de justifier leurs réponses.
- **Visionnage de la capsule (2 passes, 5 min × 2)** :
  1. Première écoute globale, sans interruption.
  2. Deuxième écoute plus attentive, avec la consigne de se concentrer sur les chiffres, les langues citées, et les répartitions géographiques.
- **Post-écoute : réponse au questionnaire Google Forms (15 min)**  
Les étudiants ont directement accédé au questionnaire via un lien transmis sur leur e-mails. Ils ont répondu en autonomie sur leur smartphone.

### 4.3. Contenu du questionnaire

### 4.4. Procédure d'évaluation automatisée

À la fin de la séance, les réponses collectées via Google Forms ont été automatiquement exportées dans un fichier Google Sheets. Seule la **dernière question ouverte** a fait l'objet d'une **analyse automatisée** via **ChatGPT**, en s'appuyant sur un **prompt soigneusement conçu** par l'enseignante. Ce prompt avait pour objectif de produire une évaluation fiable, claire et utile pour l'étudiant.

Un extrait du prompt utilisé :

*Tu es un enseignant de FLE universitaire. Corrige cette réponse d'un étudiant à une question de compréhension orale. Évalue la pertinence de la réponse (sur 5 points), la clarté syntaxique (sur 3 points) et la richesse lexicale (sur 2 points). Fournis une reformulation améliorée de la réponse. Termine par un court commentaire général encourageant.*

Cette instruction a été intégrée à un outil automatisé (extension GPT pour Google Sheets), permettant un traitement semi-massif des copies.

### 4.5. Délai et vérification

- **Temps total de traitement** : 12 minutes pour l'ensemble des réponses ouvertes.
- **Échantillonnage de vérification** : l'enseignante a contrôlé manuellement 5 copies pour s'assurer de la cohérence et de la fiabilité des retours produits par l'IA.

### 4.6. Retour aux étudiants

Chaque étudiant a reçu, **le jour même**, son feedback personnalisé, transmis individuellement par message (via e-mail ou WhatsApp), incluant :

- Une **note sur 10** (pondérée sur les 4 questions),
- Un **commentaire personnalisé** sur sa réponse ouverte,
- Une **reformulation améliorée** de sa réponse, quand nécessaire,
- Un message d'encouragement ou un conseil de reformulation.

Cette modalité de correction et de feedback a permis une **réduction notable du délai de retour**, par rapport au groupe témoin

## 1. Comparaison des deux modalités d'évaluation

Cette section vise à confronter, de manière descriptive et factuelle, les deux modalités d'évaluation mises en œuvre lors de cette expérimentation. Il s'agit de comparer les aspects matériels, humains et pédagogiques observés lors des deux séances, en se fondant sur des critères objectifs : temps de correction, nature du feedback, accessibilité technologique, et réception globale du dispositif par les étudiants.

### 5.1. Critères comparatifs retenus

Les critères suivants ont été retenus pour établir la comparaison :

- **Temps de traitement** : durée nécessaire pour la correction et le retour aux étudiants
- **Type de feedback** : feedback immédiat, différé, global ou individualisé
- **Précision et clarté de l'évaluation** : qualité du retour pédagogique sur la production écrite
- **Faisabilité technique** : compatibilité avec les moyens disponibles dans un contexte universitaire algérien
- **Réception par les étudiants** : perception de l'utilité, de la clarté et de la réactivité du système d'évaluation

### 5.2. Résultats

#### 5.2.1. Tableau Comparatif récapitulatif

Critères	Groupe témoin (classique)	Groupe expérimental (avec IA)
<b>Temps de correction</b>	~1h30 pour 25 copies	~12 minutes pour 25 réponses ouvertes
<b>Type de feedback</b>	Oral général en séance suivante	Feedback individualisé, détaillé, immédiat
<b>Qualité du retour</b>	Simple note + correction	Note + commentaire personnalisé +



Critères	Groupe témoin (classique)	Groupe expérimental (avec IA)
<b>pédagogique</b>	collective	reformulation
<b>Moyens techniques</b>	Aucun outil spécialisé ; smartphones + Google Forms	Idem + Google Sheets + ChatGPT
<b>Accessibilité technologique</b>	Moyenne (nécessite présence de l'enseignant)	Bonne (semi-automatisée, supervision humaine partielle)
<b>Réaction des étudiants</b>	Peu d'interactions avec la correction	Meilleure réception et compréhension du feedback

### 5.2.2. Observations descriptives sur les deux modalités d'évaluation

Cette mise en parallèle des deux modalités d'évaluation met en évidence des écarts structurels notables à plusieurs niveaux du processus pédagogique. Le recours à une intelligence artificielle générative, en l'occurrence ChatGPT, dans l'analyse des productions écrites issues de l'activité de compréhension de l'oral, a permis d'enclencher un traitement automatisé des réponses, générant ainsi un gain de temps substantiel pour l'enseignante. Là où la correction manuelle nécessitait environ une heure et demie pour un groupe de 25 étudiants, l'automatisation a permis de réduire cette durée à moins d'un quart d'heure, tout en assurant l'extraction et la reformulation des réponses de manière systématique et ciblée.

En parallèle, la nature même du feedback fourni aux étudiants s'est trouvée transformée. Tandis que le groupe témoin recevait un retour oral globalisé, généralement différé à la séance suivante, le groupe expérimental a bénéficié de commentaires personnalisés, générés automatiquement et transmis immédiatement. Ces retours, rédigés sous forme de remarques explicites ou de suggestions de reformulation, ont été accessibles individuellement, favorisant ainsi une lecture autonome et un possible retour réflexif de la part des étudiants.

Il convient également de souligner que cette expérimentation n'a pas nécessité de moyens techniques avancés. Malgré l'absence d'un laboratoire de langues ou d'une salle multimédia équipée, les deux groupes ont pu accéder au contenu audiovisuel via leurs smartphones personnels et un haut-parleur portable. L'évaluation s'est déroulée à travers des outils numériques simples et gratuits, notamment Google Forms pour la saisie des réponses et Google Sheets pour l'extraction, ce qui a permis une équité technique entre les deux groupes.

Aucun dispositif particulier n'a été réservé à un seul groupe, ce qui garantit la comparabilité des conditions d'expérimentation.

Ce constat d'accessibilité matérielle constitue un élément central dans l'évaluation de la transférabilité de cette approche. Il montre qu'un contexte universitaire aux ressources limitées peut néanmoins accueillir des pratiques pédagogiques innovantes dès lors que l'accompagnement méthodologique est pensé en amont.

La suite de cette étude proposera une analyse qualitative approfondie des données issues des observations, en particulier les perceptions exprimées par les étudiants dans les deux groupes. Elle examinera également les limites rencontrées lors de la mise en œuvre du protocole ainsi que les ajustements à envisager pour une intégration durable et raisonnée de l'intelligence artificielle dans l'enseignement/apprentissage du FLE dans le contexte universitaire algérien

### **5.2.3. Lecture interprétative des impacts pédagogiques de l'IA**

L'analyse comparative effectuée à partir des deux protocoles expérimentaux met en évidence une série d'éléments différenciateurs en faveur de l'intégration de l'intelligence artificielle dans le processus d'évaluation pédagogique, notamment à travers l'outil ChatGPT. Cette approche a entraîné des transformations notables dans l'organisation, la temporalité et la qualité du dispositif mis en œuvre.

- Gain de temps considérable pour l'enseignante, permettant une disponibilité accrue pour d'autres tâches pédagogiques

Le recours à l'IA a permis d'automatiser la lecture, la compréhension et l'évaluation initiale des réponses écrites des étudiants, ce qui a significativement réduit la charge cognitive et temporelle de l'enseignante. Alors que la correction manuelle nécessitait près d'une heure et demie pour traiter l'ensemble des copies du groupe témoin, la correction assistée par IA a été réalisée en moins de quinze minutes, avec une précision syntaxique et lexicale notable. Ce gain de temps libère l'enseignante pour d'autres tâches à forte valeur ajoutée : élaboration de contenus, suivi individualisé, ou encore réflexion didactique sur les activités proposées.

- Enrichissement du retour fourni aux étudiants, qui devient plus engageant, précis et individualisé

Le feedback généré par l'IA ne se limite pas à une simple validation ou invalidation des réponses. Il peut inclure des reformulations, des suggestions de correction, voire des commentaires sur la clarté ou la pertinence du propos. Cette richesse du retour transforme l'acte d'évaluation en un moment d'apprentissage personnalisé. Contrairement au groupe

témoin qui a reçu un retour global et différé, les étudiants du groupe expérimental ont bénéficié d'un accompagnement plus direct et ciblé, susceptible de renforcer la compréhension des erreurs et de favoriser l'autonomie dans la révision.

- Réaction plus vive des apprenants, en particulier sur les reformulations et remarques faites par l'IA, perçue comme neutre et objective

Les retours produits par l'intelligence artificielle ont été perçus par plusieurs étudiants comme dénués de jugement personnel, ce qui peut réduire l'effet anxiogène associé à l'évaluation. Le fait que le commentaire provienne d'un agent non humain a généré, chez certains apprenants, un sentiment de confiance accru dans l'équité du traitement. De plus, les reformulations proposées par l'IA ont souvent été perçues comme des modèles linguistiques à imiter, incitant les étudiants à comparer leur production initiale avec la version optimisée générée automatiquement.

- Toutefois, la dépendance aux outils externes (connexion, plateformes IA) nécessite un accompagnement méthodologique préalable et un encadrement pédagogique solide

Malgré les avantages constatés, l'introduction d'une technologie basée sur l'IA ne peut se faire sans une préparation technique et pédagogique rigoureuse. L'utilisation de plateformes comme ChatGPT suppose un accès stable à Internet, une familiarité minimale avec les interfaces numériques, et un encadrement clair sur les objectifs de l'exercice. Sans ces conditions, le risque de confusion ou d'incompréhension de la part des étudiants est réel. Par ailleurs, l'enseignante doit s'assurer que l'utilisation de l'IA reste au service des apprentissages, sans substitution totale du jugement pédagogique humain.

## 6. Discussion

### 6.1. Apports perçus de l'intelligence artificielle par les étudiants

À l'issue de l'expérimentation menée auprès des deux groupes (témoin et expérimental), une enquête exploratoire a été mise en place afin d'apprécier **l'impact subjectif et pédagogique** de l'activité sur les apprenants. Cette démarche visait à confronter l'efficacité perçue des deux modalités d'évaluation (classique vs. assistée par intelligence artificielle) à travers le **regard des premiers concernés : les étudiants eux-mêmes**. L'enquête a reposé sur deux outils complémentaires :

- un **questionnaire anonyme et court**, distribué immédiatement après la séance d'évaluation, composé de questions fermées (type Likert) et ouvertes ;

- un **échange oral collectif semi-dirigé**, permettant de nuancer les réponses et d'ouvrir un espace de parole réflexive.

L'objectif était double :

- Recueillir les impressions générales des étudiants sur l'activité de compréhension de l'oral (contenu, consignes, déroulement) ;
- Évaluer leur perception des modalités d'évaluation, et plus précisément leur vécu face à la correction classique ou générée par l'IA (ChatGPT).

Chez les étudiants du groupe expérimental, plusieurs tendances récurrentes ont émergé, mettant en lumière des dynamiques intéressantes sur les plans cognitif, affectif et technopédagogique :

#### ▪6.1.1. **Appréciation de la rapidité de réponse**

Une majorité significative des participants a mis en avant le bénéfice immédiat d'un retour quasi instantané sur leur production. Là où la correction manuelle par l'enseignant nécessite souvent un délai (allant de quelques heures à plusieurs jours), l'IA a permis un traitement immédiat, directement intégré à la séance. Cette immédiateté a eu pour effet de renforcer l'attention et la mémorisation : les étudiants ont pu comprendre et corriger leurs erreurs pendant que l'activité était encore fraîche dans leur esprit. Cela s'est avéré particulièrement bénéfique pour les étudiants ayant un profil auditif ou une mémoire de travail limitée, qui tirent davantage profit d'un retour en temps réel. Cette temporalité rapprochée entre production et feedback favorise un apprentissage situé, en ancrant les corrections dans le contexte direct d'usage.

#### ▪6.1.2. **Clarté et lisibilité du feedback**

Un autre aspect fréquemment cité est la clarté du langage utilisé par l'IA. Les retours générés étaient jugés précis, bien structurés, et exempts de jargon technique inutile. Plusieurs étudiants ont déclaré avoir mieux compris les remarques formulées par l'IA que celles données par l'enseignant lors des séances précédentes. Cela tient notamment à la capacité de ChatGPT à adapter son discours, à reformuler les erreurs de manière pédagogique et à proposer des améliorations concrètes et contextualisées. L'usage de formulations neutres, positives et bienveillantes, sans jugement implicite, a également été relevé comme un point fort, réduisant l'effet « sanction » souvent associé à l'évaluation.

### **6.1.3. Curiosité accrue vis-à-vis de l'outil**

L'introduction de l'IA dans le processus pédagogique a suscité une curiosité spontanée. Plusieurs étudiants ont manifesté un intérêt marqué pour l'outil lui-même, posant des questions sur son fonctionnement, son entraînement linguistique, ses limites et ses usages potentiels hors du cadre académique. Certains ont même exprimé le souhait d'utiliser l'IA de manière autonome pour réviser leurs productions écrites ou orales, voire pour préparer leurs examens. Ce comportement témoigne d'un décloisonnement de l'usage de l'outil, qui devient non plus un simple instrument ponctuel, mais un compagnon d'apprentissage potentiel. D'un point de vue didactique, cette appropriation est précieuse : elle traduit une motivation intrinsèque et une ouverture vers une autonomie numérique raisonnée, à condition que l'usage soit accompagné.

### **▪ 6.1.4. . Neutralité perçue du jugement**

L'un des apports les plus notables concerne la dimension affective de la correction. Pour une partie des étudiants, notamment ceux ayant une faible estime de soi à l'écrit, le fait que le feedback provienne d'une entité non humaine a été vécu comme moins stressant, moins jugeant.

L'IA a été perçue comme impartiale, équitable, offrant une forme de neutralité appréciée, en particulier chez les étudiants sensibles à l'évaluation. Cette absence de subjectivité a contribué à rétablir une relation plus apaisée à l'erreur, considérée comme un élément normal du processus d'apprentissage, et non comme une faute stigmatisante. Dans un contexte universitaire algérien encore très marqué par des formes d'évaluation descendantes, parfois autoritaires, cette approche a été perçue comme renouvelant les rapports pédagogiques traditionnels, en introduisant plus d'équité, voire de dialogue.

Ces constats, bien que préliminaires, ouvrent des perspectives pédagogiques intéressantes. Ils montrent que l'usage de l'intelligence artificielle, lorsqu'il est intégré de manière didactique et accompagnée, peut non seulement améliorer l'efficacité du feedback, mais aussi transformer positivement le rapport des apprenants à l'évaluation et à l'erreur. Au-delà de ses fonctions techniques, l'IA semble avoir joué un rôle structurant, rassurant et motivant, en facilitant la compréhension, en stimulant la curiosité et en réduisant la charge émotionnelle souvent associée à l'évaluation académique. Ces premiers résultats appellent à une exploration plus large, avec des échantillons plus variés et sur des temps plus longs, pour mieux cerner la manière dont l'intelligence artificielle

pourrait durablement s'inscrire dans les pratiques pédagogiques universitaires, notamment en FLE.

Malgré l'accueil globalement favorable de l'outil d'intelligence artificielle, certains étudiants du groupe expérimental ont également exprimé des réserves critiques, qui méritent d'être prises en compte dans une optique d'amélioration des pratiques.

Un premier point soulevé concerne la capacité de l'IA à interpréter les intentions implicites, en particulier dans des productions orales ou écrites marquées par des références culturelles, idiomatiques ou pragmatiques. Pour certains apprenants, l'outil semble limité dans sa compréhension fine des contextes d'énonciation spécifiques, notamment lorsqu'ils mobilisent des tournures expressives propres à leur culture, ou des implicites liés à l'humour, à l'ironie ou à des références locales. Cette limite perçue rejoint certaines critiques faites aux IA génératives dans la littérature scientifique, où l'on souligne leur tendance à appliquer des normes générales sans accès à la contextualisation sociolinguistique fine (Clément, 2023 ; Garcia & Lahire, 2022).

Par ailleurs, plusieurs étudiants ont regretté l'absence d'explication orale ou de reformulation humaine. Si les retours écrits générés par l'IA étaient globalement jugés clairs, ils ne permettaient pas toujours une négociation du sens ou une reformulation en cas d'incompréhension. Certains participants ont souligné qu'un enseignant humain, dans une situation de correction en présentiel, aurait pu reformuler une consigne, adapter le niveau de langue, ou encore dialoguer avec l'étudiant pour ajuster la correction à son niveau ou à son intention. L'absence de cet échange dialogique a parfois été perçue comme un frein à la pleine assimilation du feedback, en particulier chez les étudiants ayant un besoin d'explication orale ou de validation affective.

Ces remarques mettent en lumière une tension importante dans l'intégration de l'intelligence artificielle : si l'IA peut efficacement compléter le travail enseignant, elle ne saurait s'y substituer entièrement, notamment dans des contextes d'apprentissage où la dimension humaine, relationnelle et socio-affective joue un rôle central. Elles appellent à penser l'IA non pas comme un outil autonome, mais comme un assistant pédagogique à articuler avec la médiation humaine, dans une logique de complémentarité.

## 6.2. Limites de l'expérimentation

Si les résultats de cette expérimentation sont encourageants et permettent d'ouvrir des pistes prometteuses quant à l'intégration de l'intelligence artificielle dans l'enseignement/apprentissage du FLE, il est toutefois nécessaire de prendre en compte un certain nombre de limites méthodologiques, qui nuancent la portée des conclusions.

- **Durée courte de l'intervention** L'expérimentation s'est déroulée sur une séance unique dans chacun des deux groupes observés. Cette brièveté limite considérablement l'analyse de l'impact réel de l'IA sur les apprentissages à moyen ou long terme. L'absence de suivi post-activité ou d'évaluation différée ne permet pas de mesurer l'évolution des compétences, ni la consolidation des acquis dans le temps. Une implantation sur plusieurs séances, voire un semestre complet, serait nécessaire pour évaluer la progression effective des étudiants et identifier d'éventuels effets cumulatifs ou régressifs.
- **Effectif restreint et non diversifié** Le nombre de participants correspondait à la taille habituelle d'un groupe de licence (environ 25 étudiants), ce qui garantit une validité écologique à l'échelle de la classe. Toutefois, cet échantillon reste statistiquement limité pour une généralisation des résultats à une population plus large. De plus, tous les étudiants appartenaient à un même département universitaire, partageant des profils linguistiques et culturels proches. Il serait donc pertinent d'élargir l'échantillon à d'autres contextes (niveaux, régions, universités) afin de croiser les variables et de confirmer ou d'infirmer les tendances observées.
- **Risques de biais expérimental** L'introduction de l'intelligence artificielle dans un cadre pédagogique encore peu familier pour les étudiants a pu susciter un effet de nouveauté, susceptible de biaiser leur perception. La curiosité initiale ou l'enthousiasme technologique peuvent avoir amplifié les réactions positives, indépendamment de la réelle efficacité du dispositif. Par ailleurs, la conscience de participer à une expérimentation (effet Hawthorne) peut avoir modifié le comportement des étudiants, qui auraient agi différemment dans un cadre totalement naturel ou routinier. Ces éléments invitent à relativiser certaines interprétations, et soulignent la nécessité d'intégrer des phases d'expérimentation plus longues, intégrées au dispositif habituel d'enseignement.

### 6.3. Conditions nécessaires à une intégration réussie de l'IA

Les observations issues de cette expérimentation, croisées aux retours recueillis auprès des étudiants et de l'enseignante, permettent de dégager plusieurs conditions fondamentales pour une intégration réussie de l'intelligence artificielle (IA) dans le cadre universitaire algérien. L'enjeu ne réside pas uniquement dans l'usage ponctuel de ces outils, mais dans leur articulation cohérente avec les objectifs pédagogiques, les réalités institutionnelles et les compétences numériques des apprenants.

- Accessibilité des outils numériques :

Un des points positifs relevés au cours de cette expérimentation est la possibilité d'utiliser des outils gratuits, légers et peu gourmands en ressources, tels que ChatGPT en version web. Cette accessibilité permet une démocratisation de l'innovation pédagogique, même dans des contextes contraints matériellement, comme l'absence de salle multimédia ou de postes informatiques individuels. L'usage de smartphones personnels, associés à une connexion mobile ou partagée, constitue un levier pragmatique pour introduire progressivement l'IA dans les pratiques de classe, à condition que l'activité reste simple, ciblée, et bien encadrée.

- Accompagnement méthodologique et pédagogique :

L'efficacité de l'IA ne réside pas uniquement dans la puissance de ses algorithmes, mais dans la manière dont elle est intégrée dans la dynamique d'enseignement/apprentissage. Il est donc essentiel que l'introduction d'un outil comme ChatGPT s'accompagne d'un temps de formation ou de sensibilisation, même minimal, à destination des étudiants. Cela inclut des indications sur les usages possibles, les limites de l'outil, la formulation des requêtes, ou encore l'interprétation des réponses. De plus, l'activité doit être pensée dans un cadre pédagogique structuré, avec des objectifs clairs, des consignes précises et un rôle explicite attribué à l'IA (aide à la correction, enrichissement lexical, reformulation, etc.). Sans cet accompagnement, l'usage risque de devenir soit inefficace, soit détourné de ses finalités initiales.

- Posture éthique et regard critique :

Enfin, l'un des enjeux majeurs de l'intégration de l'intelligence artificielle dans la formation universitaire concerne le développement d'un esprit critique chez les apprenants. L'IA ne doit en aucun cas être perçue comme une source de vérité absolue, mais comme un outil d'assistance linguistique parmi d'autres. Il est nécessaire d'enseigner aux étudiants à évaluer la pertinence, la fiabilité et les limites des réponses générées par l'IA. Cette posture critique



est indispensable pour éviter les risques d'automatisme, de dépendance ou d'erreurs non repérées. Elle contribue également à former des usagers responsables et autonomes, capables de tirer parti de la technologie sans se laisser gouverner par elle.

## **Conclusion**

L'expérimentation menée, bien que circonscrite à une séance unique et à un effectif limité, offre des enseignements précieux sur l'introduction de l'intelligence artificielle dans un contexte universitaire algérien, en particulier dans le cadre de l'évaluation de la compréhension de l'oral en FLE. Les résultats montrent une réception globalement positive de l'outil par les étudiants, qui ont apprécié la rapidité, la clarté et la neutralité des retours fournis par l'IA. Ces éléments semblent favoriser à la fois l'engagement cognitif et affectif des apprenants, et ouvrent la voie à de nouvelles formes d'interaction pédagogique.

Cependant, l'analyse met également en lumière certaines limites méthodologiques (durée restreinte, biais de nouveauté, effectif réduit), ainsi que des réserves pédagogiques légitimes soulevées par les étudiants, notamment sur le manque d'explication orale ou de prise en compte des nuances culturelles dans les réponses de l'IA. Ces éléments soulignent la nécessité d'un encadrement rigoureux, fondé sur des choix didactiques réfléchis et sur une sensibilisation critique à l'usage de ces outils.

En somme, cette étude exploratoire démontre que l'IA peut constituer un levier efficace de soutien à l'apprentissage, à condition qu'elle soit utilisée comme un complément intelligent et non comme un substitut à l'enseignant. L'avenir de ces dispositifs dépendra largement de la capacité des institutions à intégrer ces technologies de manière éthique, structurée et accessible, tout en renforçant la formation des enseignants et des étudiants à une utilisation critique et éclairée.

## **CONCLUSION GENERALE**

Ce mémoire s'est inscrit dans une réflexion contemporaine essentielle autour de l'intégration des technologies numériques dans l'enseignement des langues, plus précisément dans l'apprentissage du français langue étrangère (FLE) dans un contexte universitaire algérien. L'objectif principal était d'évaluer la pertinence et l'impact d'un outil d'intelligence artificielle, ChatGPT, dans une activité pédagogique ciblée portant sur la compréhension de l'oral et l'évaluation des productions écrites. À travers une expérimentation comparative entre un groupe témoin, évalué manuellement, et un groupe expérimental, bénéficiant d'une correction assistée par IA, cette étude a permis de mettre en lumière plusieurs éléments clés liés à l'usage de ces nouvelles technologies dans un cadre éducatif.

L'un des premiers enseignements majeurs concerne le gain de temps considérable pour l'enseignante, rendu possible par l'automatisation des corrections et la rapidité de génération des feedbacks. Alors que la correction manuelle des travaux nécessitait un investissement temporel important, qui réduisait la disponibilité pour d'autres tâches pédagogiques essentielles (préparation de cours, accompagnement personnalisé, suivi des étudiants), l'utilisation de l'intelligence artificielle a permis de diminuer drastiquement ce temps. Cette optimisation offre une marge de manœuvre précieuse aux enseignants, favorisant une meilleure gestion de la charge de travail et une réallocation des efforts vers des activités à plus forte valeur ajoutée pédagogique.

Par ailleurs, l'expérimentation a montré que le retour fourni par l'IA est non seulement plus rapide mais également plus riche et individualisé. Contrairement à un feedback global et souvent différé, l'évaluation assistée par ChatGPT permet de cibler précisément les erreurs lexicales, grammaticales ou syntaxiques, et de formuler des remarques claires et accessibles. Cette personnalisation du retour contribue à une meilleure compréhension des erreurs par les apprenants, qui peuvent ainsi corriger leurs productions dans la continuité immédiate de l'activité. Ce phénomène, loin d'être anodin, favorise un processus d'apprentissage actif et réflexif, où l'étudiant est invité à s'impliquer dans sa propre progression.

Un autre aspect positif relevé est la réaction plus vive des étudiants face à ce mode d'évaluation innovant. L'étude qualitative menée a souligné une motivation accrue, une curiosité technologique, ainsi qu'un sentiment de confiance renforcé vis-à-vis du jugement posé par l'intelligence artificielle. Le caractère neutre et impersonnel de l'IA a été perçu comme un avantage, notamment pour les apprenants plus réservés ou en difficulté, pour qui la correction humaine peut parfois générer un stress ou une appréhension. Cette dimension

psychologique, souvent négligée, est en réalité un levier important pour instaurer un climat d'apprentissage positif et encourager l'expérimentation linguistique sans crainte du jugement.

Cependant, cette intégration de l'IA ne peut être envisagée sans prendre en compte plusieurs conditions essentielles pour garantir sa pertinence et son efficacité dans le contexte spécifique de l'université algérienne. En premier lieu, l'accessibilité des outils numériques est un facteur fondamental. L'absence de salles multimédia équipées ou d'infrastructures sophistiquées ne doit pas constituer un obstacle majeur. L'utilisation d'applications web gratuites, accessibles via des smartphones ou des ordinateurs personnels, constitue une solution pragmatique permettant d'ouvrir la voie à une démocratisation progressive des technologies numériques dans les espaces pédagogiques. Cette approche garantit une équité d'accès et limite les inégalités liées aux ressources matérielles.

En second lieu, un accompagnement méthodologique solide est indispensable. Les étudiants doivent bénéficier d'une formation initiale pour comprendre le fonctionnement de l'intelligence artificielle, ses possibilités, mais aussi ses limites. L'enseignant joue ici un rôle pivot en encadrant l'usage de l'outil, en guidant la réflexion critique des apprenants, et en assurant une médiation pédagogique adaptée. Cette posture d'accompagnement favorise une appropriation réfléchie de la technologie, évitant les usages mécaniques ou passifs, et encourageant au contraire un apprentissage autonome et responsable.

Troisièmement, la dimension éthique et critique doit être systématiquement intégrée. L'intelligence artificielle ne doit pas être perçue comme une vérité absolue, mais comme une aide à la compréhension et à l'analyse. Les réponses générées par ces outils peuvent comporter des erreurs, des approximations ou des biais liés à leurs bases de données et à leurs algorithmes. Ainsi, il est crucial de développer chez les étudiants une posture critique, leur permettant de confronter les propositions de l'IA à d'autres sources, à leur propre réflexion, et à l'intervention humaine. Cette vigilance éthique est également nécessaire pour préserver la dimension humaine de l'enseignement, en évitant que la technologie ne se substitue aux interactions sociales fondamentales dans l'acte pédagogique.

Au-delà des résultats quantitatifs et qualitatifs de cette étude, il convient d'insister sur les limites méthodologiques identifiées, qui appellent à une prudence dans la généralisation des conclusions. L'expérimentation s'est déroulée sur une période relativement courte, avec une seule séance par groupe, ce qui ne permet pas d'évaluer l'évolution réelle des apprentissages sur le long terme. De plus, la taille réduite des groupes restreint la représentativité des

données. Enfin, il est possible que l'effet de nouveauté et la curiosité suscitée par l'outil aient influencé les réponses des étudiants, introduisant un biais expérimental à considérer dans l'interprétation des résultats.

Ces limites ouvrent néanmoins la voie à des perspectives de recherche prometteuses. Il serait intéressant de conduire des expérimentations plus étendues dans le temps, impliquant des cohortes plus larges, et intégrant des mesures répétées pour observer les dynamiques d'apprentissage et les effets durables de l'intégration de l'IA. Par ailleurs, des travaux complémentaires pourraient explorer l'usage de ces outils dans d'autres compétences linguistiques, comme la production orale, la compréhension écrite, ou la médiation interculturelle. Ces études contribueraient à mieux définir les cadres d'utilisation optimaux et à adapter les outils aux besoins spécifiques des apprenants algériens.

Enfin, sur le plan professionnel et institutionnel, ce mémoire souligne l'importance d'accompagner la modernisation des formations universitaires par une mise à jour des compétences des enseignants, un investissement dans des infrastructures numériques adaptées, et une réflexion collective sur les pratiques pédagogiques. L'introduction de l'intelligence artificielle dans l'enseignement ne doit pas être isolée, mais s'inscrire dans une démarche globale d'innovation pédagogique, conjuguant tradition et modernité, rigueur scientifique et ouverture technologique.

En conclusion, cette recherche apporte un éclairage précieux sur les potentialités offertes par l'intelligence artificielle pour enrichir l'enseignement du FLE en Algérie. L'IA, loin de remplacer l'enseignant, s'affirme comme un partenaire puissant et complémentaire, capable de transformer les pratiques d'évaluation et de feedback. À condition d'être intégrée avec discernement, accompagnée méthodologiquement et placée sous une posture éthique, elle ouvre de nouvelles voies pour répondre aux défis de la formation linguistique dans un monde numérique en constante évolution.

## **REFERENCES BIBLIOGRAPHIQUES**

## 1. Ouvrages

- Alderson, J. C., Clapham, C., & Wall, D. (1995). *Language test construction and evaluation*. Cambridge University Press.
- Beacco, J.-C. (2007). *Didactique des langues et construction européenne*. Didier.
- Berthet, A. (2019). *Évaluation en FLE : Défis et enjeux pédagogiques*. Presses Universitaires de France.
- Bessonnat, A. (2022). *Pédagogie et intelligence artificielle : vers un nouvel équilibre*. Armand Colin.
- Bouvier, A., & Trebbi, D. (2020). *Évaluer pour faire apprendre*. De Boeck Supérieur.
- Brookhart, S. M. (2013). *How to create and use rubrics for formative assessment and grading*. ASCD.
- Cassany, D. (2012). *L'évaluation des compétences en langues étrangères*. Éditions Retz.
- Collins, A., & Hull, A. (2017). *L'IA et l'apprentissage personnalisé : Quels enjeux pour l'éducation ?* Presses Universitaires de France.
- Ducrot, O. (2004). *Le feedback en didactique des langues : Approches théoriques et pratiques*. Hachette.
- Goh, C. (2017). *La production orale en FLE : Stratégies d'évaluation et d'enseignement*. CLE International.
- Lemaire, L. (2015). *L'évaluation en langue étrangère : Théories et pratiques*. Nathan.
- Lemoine, D. (2019). *L'intelligence artificielle : Hype ou réalité pour l'éducation ?* Éditions CNRS.
- Lemoine, P. (2019). *Les enjeux éthiques de l'intelligence artificielle à l'école*. Éditions du CNRS.
- Legros, S. (2021). *L'autoévaluation en contexte numérique : Enjeux et pratiques*. Éditions L'Harmattan.
- Meunier, F. (2019). *Évaluation formative et développement des compétences langagières*. Hachette FLE.

- Meyer, C. (2020). *L'évaluation en contexte plurilingue et interculturel*. Presses Universitaires de Grenoble.
- Mialaret, G. (2018). *L'intelligence artificielle et son impact sur l'éducation*. Éditions L'Harmattan.
- Mialaret, J.-M. (2018). *Apprentissage, IA et évaluation adaptative*. Presses Universitaires de France.
- Puren, C. (2016). *L'évaluation des compétences langagières en FLE*. Didier.
- Rivière, M. (2017). *La didactique des langues et l'évaluation des compétences*. De Boeck.
- Schön, D. A. (1983). *The reflective practitioner: How professionals think in action*. Basic Books.
- Vaillancourt, L. (2018). *La motivation des apprenants dans le cadre de l'évaluation en FLE*. Presses de l'Université du Québec.

## **2. Articles scientifiques et périodiques**

- (Articles publiés dans des revues)
- Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*.
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5–31.
- Byram, M. (2000). Assessing intercultural competence in language learning. *Sprogforum*, 18(6), 8–13.
- Popenici, S. A. D., & Kerr, S. (2017). Exploring the impact of artificial intelligence on teaching and learning in higher education. *Research and Practice in Technology Enhanced Learning*, 12(1), 1–13.
- Stiggins, R. (2005). From formative assessment to assessment for learning: A path to success in standards-based schools. *Phi Delta Kappan*, 87(4), 324–328



### **3. Sitographie**

- TV5MONDE. (s.d.). *Infographie : les langues du monde* [Vidéo]. TV5MONDE.  
<https://www.tv5monde.com/information/infographies/langues-du-monde>

## **ANNEXES**

## Annexe 1 – Questionnaire de compréhension orale

**Support utilisé :** Capsule TV5MONDE – *Infographie : Les langues du monde*

**Lien :** <https://www.tv5monde.com/information/infographies/langues-du-monde>

**Objectif :** Évaluer la compréhension orale d'un document authentique court (niveau B1/B2)

### Questionnaire

#### Partie A – Prévisionnage (activation des connaissances) – 2 points

1. **À votre avis, quelle est la langue la plus parlée dans le monde ?** (*QCM – 1 pt*)
  - a. L'anglais
  - b. Le mandarin chinois
  - c. L'espagnol
2. **Combien de langues estime-t-on parlées aujourd'hui dans le monde ?** (*QCM – 1 pt*)
  - a. Environ 1 000
  - b. Environ 3 000
  - c. Environ 7 000

#### Partie B – Visionnage (compréhension ciblée) – 4 points

3. **Qu'est-ce qu'une langue « en danger » selon la vidéo ?** (*Réponse courte – 1 pt*)
4. **Quelle conclusion principale propose la vidéo concernant la diversité linguistique ?** (*Réponse courte – 1 pt*)
5. **Combien de langues risquent de disparaître d'ici la fin du siècle ?** (*QCM – 1 pt*)
  - a. Environ 600
  - b. Environ 1 500
  - c. Environ 3 000
6. **Quels continents sont les plus riches en diversité linguistique ?** (*Réponse courte – 1 pt*)
7. **Quelles sont les deux familles de langues évoquées dans la vidéo ?** (*Réponse courte – 1 pt*)

#### Partie C – Post-visionnage (réaction personnelle) – 2 points

8. **Pourquoi, selon vous, est-il important de préserver les langues en voie de disparition ?** (*Question ouverte – 2 pts : jugement, formulation et pertinence*)

#### Barème total : /10 points

- Partie A : 2 pts
- Partie B : 6 pts
- Partie C : 2 pts

## Annexe 2 – Grille d'évaluation : Groupe Témoin

**Méthode** : Passation classique (écoute de la vidéo deux fois, puis réponse écrite sans assistance numérique ou IA).

Numéro de question	Critères d'évaluation	Réponse attendue	Barème
1	Identification correcte	Le chinois (mandarin)	1 pt
2	Donnée chiffrée correcte	Environ 7 000 langues	1 pt
3	Reformulation de la définition	Une langue qui risque de disparaître	1 pt
4	Compréhension de la conclusion	La diversité linguistique est en danger / à protéger	1 pt
5	Donnée chiffrée	Environ 3 000 langues menacées	1 pt
6	Réponses multiples : continents	Afrique, Asie / Océanie (1 pt chacune)	2 pts
7	Nom de familles linguistiques	Indo-européenne, sino-tibétaine	1 pt
8	Pertinence de la réponse personnelle	Argument clair, vocabulaire adapté	2 pts

### Annexe 3 – Grille d'évaluation : Groupe Expérimental

**Méthode** : Accès à des outils d'IA pour reformuler, transcrire, ou vérifier leurs réponses (ex : ChatGPT, speech-to-text, glossaire IA).

Numéro de question	Critères d'évaluation	Réponse attendue	Barème	Indications spécifiques
1	Identification + précision du dialecte (si mentionné)	Le chinois mandarin	1 pt	Reformulation tolérée
2	Réponse chiffrée proche	Environ 7 000	1 pt	Acceptation de 6 500 à 7 000
3	Explication reformulée par l'IA correctement exploitée	Langue menacée de disparition selon l'UNESCO	1 pt	Reformulation valorisée
4	Synthèse personnelle ou reformulée	Diversité linguistique en danger / à valoriser	1 pt	Formulation plus fluide attendue
5	Réponse numérique précise	Environ 3 000	1 pt	Donnée correcte + justification
6	Réponse multiple + argumentation IA	Afrique, Asie, Océanie	2 pts	Bonus si enrichi avec ex. de langues
7	Familles de langues citées + brève définition	Indo-européenne, sino-tibétaine	1 pt	Reformulation tolérée
8	Réflexion argumentée soutenue par un outil IA ou citation	Préserver les langues, patrimoine culturel, etc.	2 pts	Bonus si citation IA pertinente

### Annexe 3 Exemple de Réponse d'étudiant – Groupe Expérimental avec feedback


1. b) Le chinois mandarin
2. a) moins de 1000
3. Une langue danger cé quand les gens parle pas, alor elle vas mourir
4. je pense c'est 20%
5. les continents comme Asie et australie est plus parler des langues
6. lang indo et lang européenne
7. faut pas perdre les langue parce que c'est important pour les traditions, le gens, la vie cultureel

### Feedback individualisé – Groupe expérimental

**Cher étudiant,**

Merci pour ta participation à cette activité de compréhension orale basée sur la vidéo *Infographie : les langues du monde*.

#### **Points positifs :**

- Tu as correctement identifié la langue la plus parlée dans le monde (le chinois mandarin).
- Tu as fait un effort d'argumentation personnelle à la fin, ce qui est très appréciable.
- Tes idées sont compréhensibles malgré les erreurs.
-  **Axes d'amélioration :**
  - **Pré-visionnage :** Tu as sous-estimé le nombre de langues parlées dans le monde (moins de 1000), alors qu'il s'agit de plus de 6000.
  - **Compréhension du contenu :** Tu as mal retenu le chiffre concernant les langues menacées (50 %, pas 20 %), et la formulation de la définition d'une langue en danger est maladroite.
  - **Orthographe et grammaire :** De nombreuses fautes nuisent à la clarté (ex. : *langue danger cé, les langue vas mourir, le gens, la vie cultureel*).
  - **Lexique et précision :** Attention aux termes trop vagues ou mal orthographiés (ex. *lang indo, lang européenne*).

#### **Recommandations :**

- Revois les notions-clés sur la diversité linguistique évoquées dans la vidéo.
- Utilise des phrases complètes avec sujet-verbe-complément.
- Réécris les réponses 3, 4 et 6 en t'aidant des bons exemples et en réécoutant la vidéo.

## **Annexe 4 : Description de la vidéo utilisée**

- **Titre** : *Infographie : les langues du monde*
- **Source** : TV5MONDE – La chaîne de la francophonie
- **Durée** : environ 1 minute 50
- **Lien** : <https://www.youtube.com/watch?v=FYyNzmJrF6Q>
- **Résumé du contenu** :

Cette capsule vidéo présente un panorama visuel et chiffré de la diversité linguistique dans le monde. Elle met en avant le nombre de langues parlées actuellement, les langues les plus utilisées, la répartition géographique des langues, les familles linguistiques principales, ainsi que la menace de disparition pesant sur certaines langues. Elle se conclut par un message appelant à préserver cette richesse linguistique, qui constitue un patrimoine immatériel précieux.

## Résumé

Ce mémoire explore l'impact de l'intelligence artificielle dans l'évaluation de la compréhension de l'oral en FLE. Deux groupes d'étudiants ont été exposés à une vidéo authentique et évalués via un questionnaire. Le groupe expérimental a bénéficié d'un accompagnement numérique IA. Les résultats montrent une amélioration significative des performances dans ce groupe. L'étude ouvre des perspectives pour une intégration pédagogique des outils numériques.

**Mots-clés :** compréhension orale, FLE, intelligence artificielle, évaluation, vidéo pédagogique

## Summary

This dissertation investigates the impact of artificial intelligence in assessing oral comprehension in French as a Foreign Language (FFL). Two student groups watched an authentic video and were evaluated via a questionnaire. The experimental group received AI-based digital support. Results show significant performance improvement in this group. The study suggests promising avenues for the pedagogical integration of digital tools.

**Keywords:** listening comprehension, FFL, artificial intelligence, assessment, educational video

## الملخص

يبحث هذا البحث في تأثير الذكاء الاصطناعي على تقييم فهم السمع في تعليم الفرنسية كلغة أجنبية. تم عرض فيديو أصيل على مجموعتين من الطلبة وتم تقييمهم عبر استبيان. استفادت المجموعة التجريبية من دعم رقمي يعتمد على الذكاء الاصطناعي. أظهرت النتائج تحسناً ملحوظاً في أداء هذه المجموعة. تفتح الدراسة آفاقاً جديدة لإدماج الأدوات الرقمية في التدريس.

**الكلمات المفتاحية:** الفهم السمعي، الفرنسية كلغة أجنبية، الذكاء الاصطناعي، التقييم، الفيديو التربوي