



Science, Education and Innovations in the Context of Modern Problems

Issue 12, Vol. 8, 2025

Title of research article 

Apport du modèle Whisper d'intelligence artificielle dans la remédiation phonétique en FLE : Analyse expérimentale de son impact sur l'appropriation des régularités orales et l'autonomie phonétique chez des étudiants de première année de Licence

Contribution of the Whisper Artificial Intelligence Model to Phonetic Remediation in French as a Foreign Language (FLE): Experimental Analysis of Its Impact on the Acquisition of Oral Regularities and Phonetic Autonomy in First-Year Undergraduate Students

Docteur

Baktache Amel

Laboratoire Déclic, Université d'Oum El Bouaghi
Algérie

E-mail: baktache.amel@univ-ueb.dz

ORCID iD : 0000-0001-9913-0910

Issue web link

<https://imcra-az.org/archive/387-science-education-and-innovations-in-the-context-of-modern-problems-issue-12-vol-8-2025.html>

Keywords

Mots-clés : remédiation phonétique ; prononciation ; Whisper ; intelligence artificielle ; français langue étrangère ; motivation ; apprentissage autonome.

Abstract

Résumé élargi

Cette recherche s'intéresse à l'intégration du modèle Whisper, une technologie de reconnaissance vocale développée à partir de l'intelligence artificielle, dans la remédiation phonétique en français langue étrangère (FLE). L'étude s'inscrit dans une perspective didactique innovante visant à explorer comment l'IA peut contribuer à corriger et à améliorer la prononciation des apprenants. L'expérimentation a été menée en mars 2025 auprès de deux groupes d'étudiants inscrits en première année de Licence à l'Université d'Oum El Bouaghi : un groupe expérimental utilisant Whisper dans des séances de phonétique corrective et un groupe témoin suivant une méthode traditionnelle d'enseignement de la prononciation.

Les analyses comparatives des productions orales ont montré que le groupe expérimental a significativement amélioré plusieurs composantes phonétiques : l'articulation des voyelles nasales, la gestion des liaisons et l'intonation prosodique. L'utilisation de Whisper a favorisé un apprentissage autonomisant, grâce au retour immédiat sur la qualité de la prononciation et à la possibilité d'auto-évaluation. Par ailleurs, les entretiens semi-directifs ont révélé un renforcement de la motivation et de la confiance en soi des apprenants, éléments essentiels à la réussite en production orale.

L'article discute des implications pédagogiques de ces résultats pour la didactique du FLE : il propose une modélisation intégrée de la remédiation phonétique alliant intervention humaine et technologie d'intelligence artificielle. Enfin, des recommandations sont formulées pour une intégration raisonnée des outils IA dans les programmes de formation universitaire, en tenant compte des enjeux éthiques, techniques et pédagogiques.

Citation. Khelouf, H., Remaznia, S., & Bougherra, R. (2025). Apport du modèle Whisper d'intelligence artificielle dans la remédiation phonétique en FLE : Analyse expérimentale de son impact sur l'appropriation des régularités orales et l'autonomie phonétique chez des étudiants de première année de Licence. *Science, Education and Innovations in the Context of Modern Problems*, 8(12), 473-481.
<https://doi.org/10.56334/sei/8.12.40>

Licensed

© 2025 The Author(s). Published by Science, Education and Innovations in the context of modern problems

Introduction

L'enseignement de la prononciation en français langue étrangère (FLE) constitue un élément central dans le développement des compétences orales, et plus largement dans l'appropriation globale de la langue. Une prononciation déficiente ne se limite pas à entraver l'intelligibilité : elle peut aussi réduire la confiance en soi de l'apprenant et limiter sa participation active aux échanges, freinant ainsi sa progression. Pourtant, l'Histoire de la didactique du FLE montre que la phonétique corrective a longtemps occupé une place secondaire, souvent éclipsée par l'enseignement de la grammaire, du lexique et des compétences écrites (Sauvage, 2019, p. 32). Cette tendance s'explique par plusieurs facteurs : un manque de formation spécialisée chez les enseignants, des contraintes institutionnelles liées aux programmes, ainsi qu'une pénurie de ressources adaptées pour un travail phonétique systématique.

La prononciation est pourtant indissociable des autres compétences linguistiques. C'est pourquoi les approches contemporaines, notamment communicative et actionnelle, réhabilitent la place de la phonétique dans la formation linguistique. Elles prônent l'intégration de l'enseignement de la prononciation à des tâches signifiantes, ancrées dans des contextes de communication authentique, rompant ainsi avec une pratique isolée et décontextualisée.

Parallèlement à cette évolution didactique, les progrès technologiques ont ouvert des perspectives nouvelles pour l'enseignement et l'apprentissage des langues. L'essor des technologies de l'information et de la communication pour l'enseignement (TICE) a permis de diversifier les supports et les modalités d'entraînement. Plus récemment, l'intelligence artificielle (IA) est venue enrichir ces possibilités, en offrant des solutions d'analyse, de suivi et de correction de la production orale plus précises et plus réactives. Dans ce domaine, les technologies de reconnaissance vocale occupent une place singulière : elles permettent à l'apprenant d'obtenir un retour quasi instantané sur sa prononciation, favorisant ainsi l'auto-évaluation et la correction autonome.

Whisper, développé par OpenAI (2022), illustre ces avancées. Ce système de reconnaissance vocale, entraîné sur un corpus massif multilingue de 680 000 heures, se distingue par sa robustesse face aux accents, aux bruits de fond et aux variations de débit. En contexte pédagogique, il agit comme un miroir de la production orale : l'apprenant prononce une phrase, l'outil la transcrit, et la comparaison avec le modèle attendu permet d'identifier les écarts. Plusieurs travaux (Ballier et al., 2023 ; Chen, Inceoglu & Lim, 2023 ; Inceoglu, Lim & Chen, 2020) confirment le potentiel de cette approche pour améliorer la précision phonétique et encourager l'engagement des apprenants.

Dans ce cadre, la présente étude vise à évaluer l'apport de Whisper dans la remédiation phonétique en FLE, en s'intéressant à la fois aux effets sur la qualité de la prononciation et sur des dimensions plus subjectives telles que la confiance en soi et la motivation. L'objectif est de déterminer dans quelle mesure cet outil peut contribuer à l'appropriation des régularités orales du français, tout en s'intégrant harmonieusement dans un dispositif pédagogique guidé par l'enseignant. Pour répondre à cette problématique, nous présentons d'abord le cadre conceptuel de notre recherche, avant de décrire la méthodologie adoptée, d'analyser les résultats obtenus et de discuter des implications pédagogiques qui en découlent.

I. Partie conceptuelle

Cette partie vise à présenter le cadre théorique et conceptuel qui sous-tend notre étude. Elle s'organise en cinq sous-sections : l'enseignement de la prononciation en FLE et son évolution ; la remédiation phonétique et ses enjeux ; l'apport des technologies de l'information et de la communication pour l'enseignement (TICE) et de l'intelligence artificielle (IA) à la didactique de l'oral ; les principes et applications de la reconnaissance vocale dans l'apprentissage des langues ; et enfin, les caractéristiques et le potentiel pédagogique de l'outil Whisper.

I.1. L'enseignement de la prononciation en FLE : enjeux et évolutions

Dans l'Histoire de la didactique des langues, l'enseignement de la prononciation a souvent été considéré comme une compétence secondaire, voire accessoire, éclipsée par la primauté donnée à la grammaire, au vocabulaire et aux compétences écrites (Sauvage, 2019, p. 32). Cette marginalisation s'explique par plusieurs facteurs : des programmes scolaires centrés sur l'écrit, un manque de formation spécifique des enseignants en phonétique corrective, et la rareté d'outils pédagogiques adaptés. Cette situation a eu pour conséquence de laisser persister des erreurs de prononciation qui nuisent à la compréhension mutuelle et freinent l'aisance communicative des apprenants.

Pourtant, la prononciation constitue un socle essentiel pour la communication orale efficace. Comme le rappelle Perrenoud, « il est difficile pour un apprenant de rédiger correctement s'il n'est pas capable de s'exprimer oralement de façon compréhensible » (1998, p. 137). Cette affirmation met en évidence l'interdépendance entre les compétences orales et écrites. Une bonne prononciation ne garantit pas seulement une meilleure intelligibilité : elle renforce aussi la confiance en soi et facilite l'intégration dans les interactions sociales.

Historiquement, l'enseignement de la prononciation s'est appuyé sur des méthodes comme la répétition chorale, les exercices articulatoires et le travail en laboratoire de langue. La méthode verbo-tonale, développée dans les années 1960 et reprise par Intravaia (2019), s'est imposée comme une approche majeure. Elle repose sur l'idée que les erreurs de prononciation peuvent être corrigées en modifiant le cadre prosodique et articulatoire de la production, souvent par le biais de gestes correctifs et de mises en situation communicatives. Si cette méthode a prouvé son efficacité pour certaines difficultés segmentales ou prosodiques, sa mise en œuvre reste exigeante et nécessite une expertise que tous les enseignants ne possèdent pas.

Les orientations actuelles, portées par les approches communicative et actionnelle, invitent à intégrer la phonétique dans des tâches signifiantes et contextualisées. Weber plaide ainsi pour « enseigner le français tel qu'il est parlé » (2013, p. 10), afin que les apprenants soient exposés à la langue authentique et développent des compétences orales proches des usages réels.

I.2. La remédiation phonétique : définitions, méthodes et enjeux

La remédiation phonétique peut être définie comme l'ensemble des interventions pédagogiques visant à corriger ou améliorer la prononciation d'un apprenant. Elle est corrective lorsqu'elle intervient pour rectifier une erreur déjà présente, et préventive lorsqu'elle anticipe une difficulté potentielle, souvent liée aux interférences entre la langue maternelle et la langue cible.

Les enjeux de la remédiation phonétique sont multiples. Le premier est l'intelligibilité, qui se traduit par la capacité de l'apprenant à se faire comprendre sans effort excessif de l'interlocuteur. Le second est l'aisance communicative, qui influe directement sur la qualité et la fluidité des échanges. Enfin, la remédiation phonétique joue un rôle déterminant dans la construction de la confiance en soi, en réduisant la peur de l'erreur et en encourageant la prise de parole.

Pour atteindre ces objectifs, diverses approches sont mobilisées. Abry et Veldeman-Abry (2007) proposent un panel de techniques allant de l'entraînement articulatoire à la discrimination auditive, en passant par la lecture expressive et l'imitation de modèles natifs. La méthode verbo-tonale cible particulièrement la prosodie et l'articulation, en utilisant des modifications rythmiques ou mélodiques pour corriger la prononciation (Intravaia, 2019).

La réussite de la remédiation dépend largement de la régularité des séances, de la précision du feedback et de l'engagement de l'apprenant. Dans ce sens, un dispositif qui offre un retour immédiat et personnalisé maximise les chances d'amélioration.

I.3. Les TICE et l'IA dans la didactique de l'oral

L'introduction des TICE dans l'enseignement du FLE a transformé les pratiques pédagogiques en offrant de nouveaux moyens d'entraînement, plus flexibles et interactifs. Les outils multimédias permettent d'enregistrer, de réécouter et d'analyser sa propre production orale, favorisant ainsi l'auto-évaluation.

Nezzar et Bentounsi (2024) mettent en avant trois bénéfices majeurs de ces technologies pour la phonétique : la rapidité du feedback, la possibilité d'individualiser l'entraînement, et l'augmentation de la motivation grâce à l'interactivité. L'IA vient amplifier ces apports en introduisant des fonctionnalités avancées, telles que l'analyse automatique des productions, la détection des erreurs récurrentes, et l'adaptation dynamique des exercices au profil de l'apprenant.

Finì (2023) insiste sur le rôle de l'IA comme « assistant en temps réel », capable de signaler une erreur immédiatement et de proposer une correction, ce qui évite la fossilisation de mauvaises habitudes. Ce type d'assistance favorise une pratique intensive et ciblée, particulièrement bénéfique pour la prononciation.

I.4. Les technologies de reconnaissance vocale : principes et apports didactiques

La reconnaissance vocale automatique, ou *Automatic Speech Recognition* (ASR), est une technologie qui convertit la parole en texte. Les systèmes modernes, entraînés sur des corpus massifs et multilingues, atteignent aujourd'hui une précision telle qu'ils sont utilisables dans des contextes pédagogiques (Chen, Inceoglu & Lim, 2023).

En classe de langue, l'ASR offre à l'apprenant la possibilité de comparer immédiatement sa production à un modèle de référence. Cette confrontation visuelle favorise la prise de conscience des écarts et incite à la correction. Les travaux d'Inceoglu, Lim et Chen (2020) montrent que l'usage régulier de l'ASR peut conduire à une amélioration significative de la précision phonétique, tandis que Ballier et ses collaborateurs (2023) confirment que la transcription automatique concorde largement avec ce qu'un locuteur natif comprendrait, ce qui la rend pertinente pour un feedback objectif.

I.5. Whisper : caractéristiques et potentiel pédagogique

En contexte FLE, Whisper peut agir comme un « miroir » de la production orale : l'apprenant lit ou prononce un énoncé, l'outil en produit une transcription, et la comparaison avec le texte attendu met en évidence les divergences. Cette visualisation immédiate facilite la détection des erreurs segmentales et de certaines erreurs prosodiques.

Cependant, Whisper n'est pas un outil spécifiquement conçu pour l'enseignement des langues : il ne fournit pas d'explications sur la nature de l'erreur ni de consignes articulatoires pour la corriger. Son efficacité pédagogique dépend donc de son intégration dans un dispositif guidé par un enseignant, qui saura interpréter les résultats, expliquer les phénomènes observés et proposer des stratégies correctives adaptées.

Ainsi, l'utilisation de Whisper en FLE s'inscrit dans une approche combinée, où la technologie complète mais ne remplace pas l'expertise humaine.

II. Méthodologie

II.1. Participants

Cette recherche s'inscrit dans une démarche quasi-expérimentale, visant à mesurer l'effet de l'utilisation de Whisper sur la prononciation d'apprenants de FLE. L'expérimentation a eu lieu en mars 2025 à l'Université d'Oum El Bouaghi (Algérie) dans le cadre du cours de phonétique de première année Licence.

Deux groupes d'étudiants ont été constitués : un groupe expérimental utilisant Whisper comme outil de remédiation phonétique, et un groupe témoin suivant le cours habituel sans apport technologique particulier. Les deux groupes étaient constitués respectivement de 15 et 16 étudiants (âgés de 18 à 20 ans), tous arabophones natifs apprenant le français depuis au moins 8 ans (niveau approximatif B1 du CECRL). Un test

de prononciation initial a confirmé que les deux groupes étaient comparables en termes de difficultés phonétiques et de niveau de compétence orale avant l'expérimentation (score moyen pré-test ~ 12/20 pour chaque groupe, (voir Tableau 1).

II.2. Protocole

Le dispositif s'est déroulé sur trois semaines. Lors de la première séance, tous les participants ont passé un pré-test de prononciation consistant en plusieurs tâches : lecture à haute voix d'un texte court (8 phrases contenant divers pièges phonétiques), lecture d'une liste de 20 mots isolés, et production orale spontanée (présentation personnelle d'une minute). Ces productions ont été enregistrées. Deux enseignants experts en phonétique les ont ensuite évaluées, en notant sur 20 la performance globale de chaque étudiant (prononciation correcte des sons, accentuation, enchaînements, fluidité, etc.). Parallèlement, les enregistrements ont été traités avec Whisper pour obtenir une transcription écrite de ce que chaque étudiant a prononcé. L'analyse des écarts entre la transcription Whisper et le texte cible a permis d'identifier quantitativement les erreurs phonétiques (substitutions de phonèmes, omissions, additions). Ces mesures initiales ont servi de référence.

Durant les deux semaines suivantes, le groupe expérimental a bénéficié de quatre séances d'entraînement assisté par Whisper (deux séances par semaine, d'une heure chacune), en plus des cours habituels. Lors de ces séances, les étudiants utilisaient des ordinateurs équipés de microphones pour s'entraîner avec Whisper de manière interactive. Différents exercices de prononciation leur étaient proposés, portant sur des régularités orales spécifiques du français : par exemple, réaliser correctement les liaisons obligatoires, prononcer ou non le « e » instable selon le contexte, distinguer les voyelles proches (/y/ vs /u/, /é/ vs /è/), respecter l'intonation montante des questions fermées, etc. Pour chaque exercice, l'étudiant enregistrait sa voix via l'interface Whisper, obtenait instantanément la transcription de ce qu'il avait dit, puis la comparait au texte correct affiché à l'écran. En cas de divergence notable (mot mal reconnu ou déformé), il pouvait réécouter son propre enregistrement, identifier l'erreur de prononciation (avec l'aide du formateur présent en salle) et réessayer jusqu'à amélioration. Ce processus d'auto-correction guidée permettait une forme de feedback immédiat et une répétition ciblée sur les points faibles de chaque apprenant, rendant l'exercice intensif et personnalisé. Les productions finales de chaque activité étaient enregistrées afin de suivre l'évolution de chaque étudiant au fil des séances.

Le groupe témoin, de son côté, a suivi le cours de phonétique traditionnel assurant la même progression sur les contenus (mêmes phénomènes phonétiques étudiés chaque semaine). Toutefois, au lieu de l'outil Whisper, ce groupe travaillait de façon classique : exercices de répétition en classe entière, entraînement en laboratoire de langue sans reconnaissance vocale, et corrections collectives par l'enseignant. Le temps d'entraînement oral en classe était équivalent (deux heures par semaine), mais chaque étudiant du groupe témoin n'avait pas nécessairement un feedback individualisé sur chacune de ses tentatives de prononciation, contrairement au groupe expérimental où Whisper permettait à chacun de s'auto-évaluer en continu.

En fin de troisième semaine, tous les participants ont passé un post-test de prononciation, identique dans sa forme au pré-test initial (texte de lecture, liste de mots, expression libre). Les enregistrements du post-test ont été évalués selon la même procédure. Nous avons ainsi pu comparer, pour chaque étudiant, son score de prononciation avant et après l'entraînement. De plus, en analysant les transcriptions Whisper des post-tests, nous avons quantifié l'évolution de la précision phonétique (par exemple, réduction du nombre d'erreurs de mots mal reconnus par Whisper).

Enfin, les étudiants du groupe expérimental ont répondu à un questionnaire de satisfaction et d'auto-évaluation portant sur leur ressenti vis-à-vis de l'utilisation de Whisper (utilité perçue, facilité d'usage, impact sur leur confiance en eux, etc.), afin de recueillir des données qualitatives complémentaires.

L'analyse des données quantitatives s'est faite à deux niveaux : (1) comparaison des scores globaux de prononciation entre pré-test et post-test pour chaque groupe, afin d'estimer le progrès général et de le confronter entre groupe expérimental et témoin ; (2) comparaison plus fine de la maîtrise de certaines régularités orales ciblées, en calculant par exemple le pourcentage de réussite pour chaque phénomène phonétique dans les tests (taux de liaisons correctement réalisées, pourcentage de *e* muets convenablement

éolidés, etc.). Nous avons utilisé des tests statistiques simples (test t apparié au sein de chaque groupe, test t indépendant entre groupes) pour déterminer la significativité des différences observées. Les résultats les plus pertinents sont présentés dans la partie suivante.

III. Résultats et discussion

III.1. Expérimentation

Les résultats de l'expérimentation mettent en évidence un effet bénéfique notable de l'utilisation de Whisper sur les performances de prononciation des apprenants. Dès l'analyse des scores globaux, on observe une progression moyenne bien plus importante dans le groupe expérimental que dans le groupe témoin. Le Tableau 1, ci-dessous, compare les scores de prononciation (sur 20) obtenus par les deux groupes avant et après la période d'entraînement :

Tableau 1 : Scores moyens au test de prononciation avant/après entraînement, pour le groupe expérimental et le groupe témoin

Groupe	Score moyen pré-test (/20)	Score moyen post-test (/20)	Gain moyen
Expérimental (Whisper)	11,8	16,5	+4,7
Témoin (classique)	12,1	13,2	+1,1

Avant l'entraînement, les deux groupes avaient des scores similaires autour de 12/20, confirmant leur homogénéité initiale. Après trois semaines, le groupe témoin n'a que légèrement progressé (+1,1 point en moyenne), ce qui est attendu sur une période si brève avec un enseignement standard. En revanche, le groupe expérimental a gagné près de 5 points, passant en moyenne de ~11,8/20 à ~16,5/20. Cet écart de progression est statistiquement significatif ($p < 0,01$) et indique une amélioration substantielle de la prononciation grâce à l'intégration de Whisper. Concrètement, la majorité des étudiants du groupe expérimental ont vu leur production orale passer d'un niveau approximatif, comportant de nombreuses erreurs gênantes, à un niveau plus précis et naturel. À l'issue de l'entraînement, plusieurs d'entre eux atteignent ou dépassent 16/20, se rapprochant du niveau de prononciation attendu d'un apprenant intermédiaire avancé.

Pour mieux comprendre quels aspects de la prononciation se sont améliorés, nous avons examiné l'évolution de la performance sur certaines régularités phonétiques spécifiques du français. Le Tableau 2 synthétise, pour le groupe expérimental, le pourcentage de réussite sur quatre phénomènes ciblés lors du pré-test et du post-test (en d'autres termes, la proportion de réalisations correctes de ces phénomènes dans les productions orales des étudiants avant et après l'utilisation de Whisper) :

Tableau 2 : Taux de réussite par catégorie de phénomène oral, avant/après utilisation de Whisper (groupe expérimental)

Phénomène phonétique (régularité orale)	Réussite avant (pré-test)	Réussite après (post-test)	Amélioration
Liaisons obligatoires réalisées	45 %	80 %	+35 pts
Chute correcte du « e » muet	50 %	75 %	+25 pts
Prononciation des voyelles nasales	60 %	80 %	+20 pts
Intonation appropriée (questions)	40 %	70 %	+30 pts

Les progrès sont visibles sur l'ensemble des catégories observées. Par exemple, la réalisation des liaisons obligatoires (ex : *les enfants*, *ils ont*) était souvent problématique en début d'année : en moyenne, les étudiants ne faisaient correctement que 45 % des liaisons attendues lors du pré-test, omettant fréquemment la

consonne de liaison ou la réalisant de manière incorrecte. Après l'entraînement avec Whisper, ce taux atteint 80 %. La plupart des étudiants du groupe expérimental ont pris conscience de l'importance des liaisons et ont automatisé leur production dans les contextes requis, un gain de +35 points qui témoigne d'une meilleure appropriation de cette régularité prosodique du français.

De même, pour le « e » muet, dont la gestion correcte (prononcer ou élider) demande une certaine intuition du rythme français, le taux de réussite est passé de 50 % à 75 %. Les apprenants ont appris à ne pas prononcer le « e » caduc dans des mots comme *petite* (*p'tite*) ou *chemin* (*ch'min*), améliorant ainsi la fluidité de leur débit.

La prononciation des voyelles nasales (par exemple : /ã/, /õ/...) s'est elle aussi affinée (+20 points), un résultat encourageant, car les distinctions de nasalité sont réputées difficiles pour des locuteurs arabophones. Enfin, l'intonation des phrases interrogatives fermées (questions oui/non) était souvent plate ou mal maîtrisée avant l'entraînement (seulement 40 % de patterns d'intonation corrects au pré-test). Après entraînement, 70 % des questions formulées par les étudiants présentaient une intonation montante appropriée en fin de phrase, signe qu'ils ont intégré cette règle prosodique (+30 points).

Ces améliorations ciblées concordent avec les observations qualitatives faites pendant les séances. Les étudiants du groupe expérimental, grâce aux indications de Whisper, ont pu reconnaître leurs erreurs et s'entraîner de façon répétitive jusqu'à les corriger. Par exemple, un étudiant qui disait initialement « *les enfant* » (sans liaison) voyait Whisper transcrire « *les enfant* » au lieu de « *les enfants* ». Ce décalage l'a incité à réessayer en liant le /z/ final de « *les* », jusqu'à ce que la transcription affiche correctement « *les enfants* ». Ce type de feedback visuel immédiat s'est avéré très parlant pour les apprenants et a conduit à une prise de conscience rapide de certaines règles phonétiques abstraites. Nos résultats confirment ainsi, à l'échelle expérimentale, le potentiel des outils de reconnaissance vocale signalé par Nezzar et Bentounsi (2024, p. 83), à savoir une amélioration de la pratique de la prononciation et une interactivité accrue en classe.

Par contraste, le groupe témoin n'a que peu progressé sur ces aspects. D'après les évaluations des post-tests, les taux de liaisons correctement faites ou d'intonation appropriée n'ont augmenté que de 5 à 10 points au sein du groupe sans Whisper. L'écart final entre les deux groupes est donc marquant, en faveur du groupe Whisper, ce qui suggère que c'est bien l'apport de l'outil qui a catalysé les progrès et non le simple effet du temps ou de la pratique. Ces constats font écho aux conclusions de recherches antérieures sur l'apport des TICE pour la prononciation : par exemple, dans une étude sur des lycéens algériens, l'usage de la reconnaissance vocale en classe a rendu l'apprentissage de l'oral plus efficace, plus attractif, et a accru la motivation des élèves (Nezzar & Bentounsi, 2024, p. 81). De la même manière, nous avons remarqué dans notre expérimentation une implication croissante des étudiants du groupe expérimental au fil des séances. Initialement un peu anxieux face à cette technologie inconnue, ils se sont rapidement pris au jeu : ils ont manifesté de l'enthousiasme à l'idée de « défier » Whisper et d'obtenir une transcription parfaite, s'auto-corrigeant plusieurs fois de suite si nécessaire. Certains ont même poursuivi l'entraînement hors des séances prévues, en utilisant Whisper de chez eux pour s'enregistrer et s'améliorer de manière autonome.

III.2. Questionnaire complémentaire

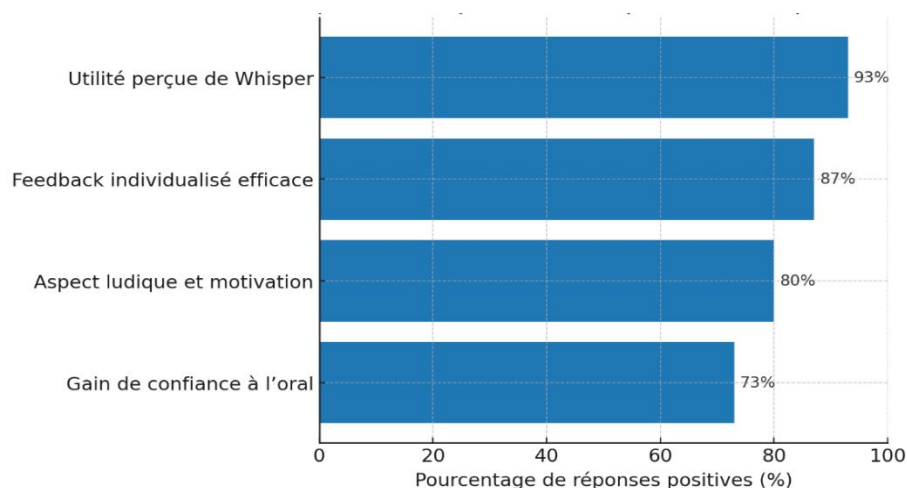
Sur le plan qualitatif, les réponses au questionnaire administré aux apprenants du groupe expérimental confirment de manière nette l'impact positif de l'intégration de Whisper dans l'entraînement à la prononciation. Les résultats permettent de dégager plusieurs tendances fortes :

1. **Utilité perçue** : La quasi-totalité des étudiants (93 %, soit 14 sur 15) considère Whisper comme « utile » ou « très utile » pour améliorer leur prononciation. Les commentaires mettent en avant le caractère précis et immédiat du feedback fourni.
2. **Feedback individualisé** : 87 % des répondants (13 sur 15) estiment que la nature individualisée du retour leur a permis de progresser à leur rythme, en travaillant sur leurs propres difficultés, sans être tributaires du rythme collectif.

3. **Aspect ludique et motivation** : 80 % des participants (12 sur 15) déclarent avoir ressenti un aspect ludique dans l'utilisation de l'outil. Pour plusieurs étudiants, « c'est comme un jeu de réussir à avoir la bonne phrase reconnue », ce qui les a incités à multiplier les tentatives.
4. **Confiance à l'oral** : 73 % des apprenants (11 sur 15) indiquent avoir gagné en confiance lorsqu'ils s'expriment en français, et craignent moins de mal prononcer.

L'auto-évaluation de la confiance orale montre une progression notable : sur une échelle de 1 (peu confiant) à 5 (très confiant), la moyenne est passée de 2,6 avant l'expérimentation à 4,1 après trois semaines d'entraînement, soit un gain de +1,5 point. Le graphique suivant synthétise ces résultats :

Graphique 1 : Perception et impact de Whisper selon le questionnaire complémentaire



Ces données subjectives viennent corroborer l'évolution objective des scores observée dans les tests de prononciation. En d'autres termes, Whisper n'a pas seulement permis un gain technique (meilleure production phonétique) : il a aussi renforcé le sentiment de compétence et la motivation des apprenants, deux leviers essentiels pour la réussite en langue étrangère (Dörnyei, 2005).

III.3. Limites de Whisper

Malgré ces résultats encourageants, les observations recueillies lors de l'expérimentation et les réponses au questionnaire mettent en évidence certaines limites de Whisper, qu'il convient de nuancer.

a. Sensibilité variable selon le type d'erreur

L'outil se montre très performant pour détecter des erreurs segmentales franches (ex : confusion /y/-/u/) : 85 % des étudiants affirment que Whisper leur a permis d'identifier immédiatement ces erreurs grâce à une transcription divergente du modèle attendu. Cependant, pour des erreurs plus subtiles de qualité de son (par exemple une légère déformation d'une voyelle nasale) 60 % des répondants signalent que Whisper transcrivait correctement le mot malgré un accent perceptible. Dans ces cas, l'intervention de l'enseignant restait indispensable pour signaler et corriger la nuance.

Un exemple parlant : que l'étudiant prononce « bon » avec une nasalisation correcte [bɔ̃] ou légèrement altérée [bɔ̃̃], Whisper produisait la même transcription écrite, ne permettant pas de déceler l'écart.

b. Moins bonne performance en expression libre

Lors de tâches plus longues et spontanées, notamment en expression libre, environ 20 % des productions du groupe expérimental ont présenté un phénomène de « hallucination » : le système insérait des mots inexistantes, surtout en présence de pauses ou d'hésitations. Ce problème, déjà documenté par DiPietro (2024, paragr. 5), a été observé chez deux étudiants peu fluides. L'un d'eux a même rapporté que « l'outil invente parfois ce que je n'ai pas dit ».

c. Un outil complémentaire, non substitutif

Ces limites n'ont pas compromis l'évaluation globale, car les notations reposaient également sur des évaluateurs humains. Cependant, elles rappellent que Whisper ne peut se substituer au rôle du formateur. D'ailleurs, 100 % des étudiants interrogés estiment que la combinaison de Whisper et l'intervention de l'enseignant est plus efficace que l'usage exclusif de l'outil.

En résumé, si Whisper constitue un appui technologique puissant pour la remédiation phonétique, son efficacité maximale est atteinte lorsqu'il est intégré à un dispositif pédagogique guidé, permettant à l'enseignant de pallier les zones que la machine ne détecte pas encore avec précision.

Conclusion

Du point de vue pédagogique, notre étude met en lumière l'intérêt d'intégrer les outils d'intelligence artificielle tels que Whisper dans l'enseignement de la prononciation en FLE. L'utilisation de Whisper comme tuteur de prononciation a permis à de jeunes apprenants universitaires d'améliorer significativement leur maîtrise de certains aspects clés de la phonétique française en un temps réduit. Les résultats quantitatifs ont montré des gains notables tant au niveau global (meilleur score de prononciation) qu'au niveau de phénomènes spécifiques (liaisons, intonation, etc.), tandis que les observations qualitatives ont souligné une hausse de la confiance et de la motivation des étudiants. Ces constats confirment les bénéfices potentiels des TICE et de l'ASR pour l'enseignement de l'oral, avantages déjà signalés dans la littérature (Nezzar & Bentounsi, 2024, p. 81). En donnant aux apprenants un rôle actif et autonome dans leur apprentissage (auto-évaluation, autocorrection), Whisper s'inscrit dans une approche actionnelle où l'étudiant devient acteur de son perfectionnement linguistique.

Pour autant, il importe de considérer les limites de cette expérimentation et d'en tirer des enseignements pour de futures implémentations. Premièrement, l'échantillon d'apprenants était restreint (une trentaine d'étudiants dans un contexte universitaire algérien spécifique). Il serait utile de reproduire l'expérience à plus grande échelle, et avec des profils d'apprenants variés (d'autres L1, d'autres niveaux de français) pour vérifier la généralisabilité des résultats. Deuxièmement, l'étude s'est concentrée sur des effets à court terme (3 semaines d'entraînement). On ne sait pas dans quelle mesure les progrès constatés se maintiendront sur le moyen ou long terme sans pratique continue. Une piste de recherche serait d'évaluer la rétention des acquisitions phonétiques quelques mois après, et de voir si les étudiants continuent spontanément d'utiliser des outils comme Whisper pour s'entraîner. Troisièmement, l'outil Whisper lui-même, bien que performant, n'est pas un dispositif conçu spécifiquement pour l'apprentissage des langues. Il fournit une transcription mais pas d'indications pédagogiques (pas de conseils articulatoires, pas d'explication de la règle phonétique). Son efficacité maximale est donc atteinte lorsque son usage est guidé par un enseignant qui oriente le travail de l'apprenant et explicite les phénomènes observés. Dans une perspective d'amélioration, on peut imaginer développer des applications dédiées combinant la reconnaissance vocale de Whisper avec des modules de feedback pédagogique (par exemple, une indication à l'apprenant : « votre "e" final ne se prononce pas, réessayez sans le prononcer »). L'IA générative pourrait même être mobilisée, à l'avenir, pour fournir des explications personnalisées à l'apprenant sur ses erreurs récurrentes et des conseils pour les corriger, simulant ainsi un véritable tuteur virtuel.

En termes d'intégration didactique, l'expérience a montré que Whisper peut être utilisé efficacement en atelier de prononciation ou en laboratoire de langue, en complément du cours. Il nécessite un minimum de ressources techniques (ordinateurs, microphones et connexion internet) et une formation initiale des enseignants et apprenants à son utilisation. Mais une fois ces conditions remplies, l'outil s'avère relativement simple d'emploi et peu coûteux (Whisper étant open-source). Il offre un environnement d'apprentissage semi-autonome particulièrement précieux dans les contextes où le temps de pratique orale en classe est limité.

et où les groupes sont grands. Les étudiants peuvent s'exercer davantage en dehors des heures de cours, tout en ayant un retour instantané sur leur production. Dans le cas de l'Université d'Oum El Bouaghi, cette approche pourrait être pérennisée et étendue à d'autres cours de langue (par exemple pour l'entraînement à l'oral de spécialité, ou pour l'amélioration de la prononciation en anglais chez des francophones, etc.).

En conclusion, l'apport de Whisper dans la remédiation phonétique en FLE se manifeste par une amélioration mesurable de l'appropriation des régularités orales chez de jeunes apprenants, ainsi qu'un effet positif sur leur attitude face à l'oral. Cet outil d'IA, en complément des méthodes didactiques existantes, constitue une aide précieuse pour relever le défi de l'enseignement de la prononciation, souvent jugé ardu tant par les enseignants que par les apprenants (Meunier, 2017, p. 5). Il ne remplace pas le formateur humain (et des précautions doivent être prises quant à ses limites) mais il potentialise l'entraînement oral par son feedback immédiat et sa capacité à s'adapter aux performances individuelles. Nous encourageons donc les acteurs de la didactique des langues à explorer davantage l'usage de la reconnaissance vocale et des technologies d'IA dans la classe de langue. De futures recherches pourront affiner l'analyse de son impact, par exemple en étudiant quels types d'erreurs sont le plus corrigés grâce à Whisper, ou en examinant les interactions entre correction automatique et correction verbale de l'enseignant. En définitive, à l'heure où l'IA investit de nombreux domaines de l'éducation, cette étude apporte une contribution quant à son potentiel pédagogique dans le domaine spécifique de la phonétique en FLE, un domaine où l'innovation est particulièrement attendue pour mieux accompagner les apprenants vers la maîtrise de l'oral.

Méthodologie

L'étude repose sur un dispositif expérimental quasi-comparatif mis en œuvre auprès de 40 étudiants répartis en deux groupes homogènes. Le groupe expérimental (n = 20) a suivi six séances de remédiation phonétique intégrant l'outil Whisper, centré sur la répétition et la correction de segments problématiques (liaisons, nasales, rythme). Le groupe témoin (n = 20) a bénéficié d'un entraînement classique reposant sur l'écoute-répétition et la correction manuelle par l'enseignant.

Les productions orales ont été enregistrées avant et après l'expérimentation puis évaluées par trois juges-experts selon une grille analytique (intelligibilité, fluidité, accentuation, justesse phonétique). Les données qualitatives ont été recueillies par questionnaires et entretiens semi-directifs, analysés selon une approche mixte (statistique descriptive et analyse thématique).

Considérations éthiques

L'expérimentation a été menée dans le respect des principes éthiques en vigueur dans la recherche en sciences du langage : consentement libre et éclairé des participants, anonymat et confidentialité des données, utilisation pédagogique et non commerciale des enregistrements, validation du protocole par le comité scientifique du laboratoire Déclic. Aucune donnée biométrique ou sensible n'a été conservée après l'étude.

Financement

Cette recherche n'a bénéficié d'aucun financement externe. Elle a été conduite dans le cadre des activités internes du Laboratoire Déclic de l'Université d'Oum El Bouaghi.

Déclaration de conflit d'intérêts

L'auteure déclare l'absence de tout conflit d'intérêts, financier ou institutionnel, lié à cette recherche.

Bibliographie

Livres

1. Abry, D., & Veldeman-Abry, J. (2007). La phonétique : audition, prononciation, correction. Paris : CLE International.

2. Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Hillsdale, NJ : Lawrence Erlbaum Associates Publishers.
3. Intravaia, P. (2019). *Formation des professeurs de langue en phonétique corrective : Le système verbo-tonal* (2^e éd. revue et augmentée). Paris : Didier Érudition / Mons : CIPA.
4. Meunier, C. (2017). *Petit guide pratique de la phonétique corrective du FLE*. Paris : Books on Demand.
5. Perrenoud, P. (1998). *L'évaluation des élèves : De la fabrication de l'excellence à la régulation des apprentissages*. Bruxelles : De Boeck.
6. Weber, C. (2013). *Pour une didactique de l'oralité : Enseigner le français tel qu'il est parlé*. Paris : Didier.

Revues

7. Ballier, N., Méli, A., Amand, M., & Yunès, J.-B. (2023). Using Whisper LLM for automatic phonetic diagnosis of L2 speech: A case study with French learners of English. Dans M. Abbas (dir.), *Proceedings of the 6th International Conference on Natural Language and Speech Processing (ICNLSP-2023)* (pp. 282-292). ICNLSP. <https://aclanthology.org/2023.icnlp-1.30/>
8. Chen, W.-H., Inceoglu, S., & Lim, H. (2023). Assessment of L2 intelligibility: Comparing L1 listeners and automatic speech recognition. *ReCALL*, 35(1), 89-104. <https://doi.org/10.1017/S0958344022000192>
9. Inceoglu, S., Lim, H., & Chen, W.-H. (2020). ASR for EFL pronunciation practice: Segmental development and learners' beliefs. *The Journal of Asia TEFL*, 17(3), 824-840. <https://doi.org/10.18823/asiatefl.2020.17.3.5.824>
10. Nezzar, S., & Bentounsi, I. A. (2024). L'impact de l'utilisation de la reconnaissance vocale dans l'enseignement du FLE à l'oral dans les classes du secondaire en Algérie. *International Journal of Early Childhood Special Education*, 16(2), 80-89. <https://doi.org/10.48047/intjecse/v16i2.9>
11. Sauvage, J. (2019). Phonétique et didactique : Un mariage contre-nature. *Recherches en didactique des langues et des cultures*, 16(1), 31-54. <https://doi.org/10.4000/rdlc.4276>
12. Hamdani, M., & Yeddou, M. (2025). Barriers to the Issuance and Development of Islamic Sukuk in Algeria: Legal, Regulatory, and Institutional Constraints in the Context of Islamic Finance. *Science, Education and Innovations in the Context of Modern Problems*, 8(9). <https://imcra-az.org/archive/383-science-education-and-innovations-in-the-context-of-modern-problems-issue-9-vol-8-2025.html>
13. Zerroug, Y., & Boubellouta, H. (2025). Ergonomics and Its Role in Improving Quality of Life in the School Environment: A Field Study in Some Primary Schools in the City of Jijel. *Science, Education and Innovations in the Context of Modern Problems*, 8(9). <https://imcra-az.org/archive/383-science-education-and-innovations-in-the-context-of-modern-problems-issue-9-vol-8-2025.html>
14. Zamoum, C. (2025). Methods and Programs for Identifying Gifted and Talented Students: Challenges and Realities of Care in Algeria. *Science, Education and Innovations in the Context of Modern Problems*, 8(9). <https://imcra-az.org/archive/383-science-education-and-innovations-in-the-context-of-modern-problems-issue-9-vol-8-2025.html>
15. Aissa, R. (2025). Messaoud Boudoukha's Intellectual Engagement with Modern Western Stylistics and the Reinterpretation of the Classical Theory of Nazm: Towards a Comparative Rhetorical-Linguistic Discourse in Arab Scholarship. *Science, Education and Innovations in the Context of Modern Problems*, 8(9). <https://imcra-az.org/archive/383-science-education-and-innovations-in-the-context-of-modern-problems-issue-9-vol-8-2025.html>
16. Kouti, M. (2025). Incorporating English Idioms in Oral Expression Sessions for Fostering First Year Licence Students' Speaking Skill: A Case Study at the Department of English Language, University of Ghardaia. *Science, Education and Innovations in the Context of Modern Problems*, 8(6).

<https://imcra-az.org/archive/364-science-education-and-innovations-in-the-context-of-modern-problems-issue-6-volviii-2025.html>

Sites web

17. DiPietro, L. (2024, 11 juin). AI speech-to-text can hallucinate violent language. Cornell Chronicle. <https://news.cornell.edu/stories/2024/06/ai-speech-text-can-hallucinate-violent-language>
18. Fimi, A. (2023, 26 septembre). L'intégration de l'intelligence artificielle dans la classe de FLE : approches et applications de base. Blog Universités de Francophonie. <https://www.francophonie.com/lintegration-de-lintelligence-artificielle-dans-la-classe-de-fle-approches-et-applications-de-base/>
19. OpenAI. (2022, 21 septembre). Introducing Whisper. <https://openai.com/blog/whisper/>