Didactique du FLES
Recherches et pratiques

Didactique du FLES
ISSN : 2826-777X

4:1 | 2025
L'intelligence artificielle générative pour l'enseignement du FLE

Conception de ressources pédagogiques pour le FLE : quelle synergie entre l'IAG, le TAL et Moodle ?

Jovan Kostov

https://www.ouvroir.fr/dfles/index.php?id=1638

DOI : 10.57086/dfles.1638

Référence électronique
Jovan Kostov, « Conception de ressources pédagogiques pour le FLE : quelle synergie entre l'IAG, le TAL et Moodle ? », *Didactique du FLES* [En ligne], 4:1 | 2025, mis en ligne le 10 juillet 2025, consulté le 26 octobre 2026. URL : https://www.ouvroir.fr/dfles/index.php?id=1638

Droits d'auteur
Licence Creative Commons - Pas d'Utilisation Commerciale - Pas de Modification 2.0 Générique (CC BY-NC-ND 4.0)

Conception de ressources pédagogiques pour le FLE : quelle synergie entre l'IAG, le TAL et Moodle ?

Jovan Kostov

PLAN

Introduction
1. État de l'art, réalités de terrain et contextualisation de l'étude
   1.1. De la génération automatique à l'intelligence artificielle générative
   1.2. L'intelligence artificielle en éducation : une affaire scientifique et politique au service des acteurs
   1.3. Qu'en est-il de la conception pédagogique ?
2. Description des outils mobilisés et du protocole expérimental
   2.1. Les outils de supervision : le CECRL, FleLex et TreeTagger
   2.2. Formats des ressources et des activités générées
   2.3. Génération des ressources textuelles et iconographiques
   2.4. Génération et adaptation des activités pédagogiques sur Moodle
3. Test du protocole et analyse des résultats
   3.1. Génération des ressources textuelles et iconographiques
   3.2. Génération et analyse des activités de compréhension, de grammaire et des ressources lexicales
      3.2.1. Activités de compréhension écrite
      3.2.2. Activités de grammaire
      3.2.3. Ressources lexicales
   3.3. Synthèse sur la génération des ressources et des activités
4. Conclusions et perspectives

TEXTE

Introduction

1 Les outils d'intelligence artificielle générative (IAG) comme ChatGPT et DALL-E, associés aux technologies de traitement automatique des langues (TAL), permettent de générer des documents en langage naturel et des images. Ces données peuvent être structurées de différentes manières afin d'être exploitées au sein de certaines plateformes éducatives, comme les *Learning Management Systems* (LMS)

de type Moodle. La synergie entre IAG et TAL ouvre ainsi plusieurs possibilités pour concevoir des ressources d'enseignement/apprentissage des langues conformes aux six niveaux du Cadre européen commun de référence pour les langues (CECRL). Le présent travail de recherche se propose d'explorer la capacité de ces outils pour produire des supports pédagogiques pour l'enseignement/apprentissage du français langue étrangère (FLE) pour les niveaux A1 et A2.

2 En effet, depuis sa démocratisation et ouverture au grand public, l'intelligence artificielle générative (IAG) a suscité un intérêt croissant dans les milieux éducatifs, notamment dans la production automatisée de contenus. Plusieurs études (Albadarin & coll., 2024 ; Kaplan-Rakowski et al., 2023 ; Nikolopoulou, 2024) ont souligné l'émergence de pratiques expérimentales visant à intégrer des outils tels que ChatGPT dans la création de ressources pédagogiques. Toutefois, ces usages restent majoritairement exploratoires et peu systématisés. Cette situation nécessite une analyse plus détaillée des outils et des conditions qui permettraient une intégration raisonnée de l'IAG dans la conception pédagogique, notamment dans les environnements d'apprentissage numériques comme Moodle. Notre travail s'inscrit dans cette dynamique, en testant une synergie entre IAG, TAL et LMS.

3 La question centrale à laquelle nous souhaitons apporter des réponses est la suivante : dans quelle mesure est-il possible de produire des ressources pédagogiques fiables et calibrées au regard des exigences du CECRL grâce à l'IAG et compatibles avec un environnement de type Moodle ? Nous formulons l'hypothèse que la combinaison entre les outils issus de l'IAG, tels que ChatGPT[1], et des outils du traitement automatique des langues (TAL), comme FleLex[2] et TreeTagger[3], permet de concevoir des ressources d'enseignement/apprentissage des langues exploitables dans des contextes pédagogiques variés et compatibles avec les plateformes numériques de type Moodle. Deux approches méthodologiques sont envisagées : une approche non supervisée reposant sur l'IAG seule (ChatGPT), et une approche supervisée, intégrant des outils du TAL et les descripteurs du CECRL pour ajuster et affiner les ressources et les activités produites. Cette recherche s'inscrit dans une perspective de recherche-développement et vise une démarche qui respecte les préconisations de sobriété numérique, alignée sur les recommandat-

ions de l'UNESCO (2019), en optimisant les outils existants pour concevoir des protocoles de conception pédagogique durables et basés sur l'existant.

1. État de l'art, réalités de terrain et contextualisation de l'étude

1.1. De la génération automatique à l'intelligence artificielle générative

4 Dans les années 1990, les chercheurs en informatique et en TAL ont produit une littérature scientifique exhaustive qui rend compte des préoccupations et des fondements du domaine de la génération automatique à l'écrit. Danlos (1991) ainsi que Zock et Sabah (1992) situent le début des travaux dans ce domaine dans les années 1970. Danlos (1991) note une certaine asymétrie historique dans le développement des domaines du TAL, considérant que la génération s'est développée bien plus tardivement que l'analyse (p. 199). Le domaine du TAL a été ponctué avec des moments clés marquant les différentes avancées comme le passage d'une approche majoritairement linguistique[4] – héritière du structuralisme et du génération automatique à une approche fondée sur des modèles statistiques (Poibeau, 2019). Selon Bouillon (1998), la génération automatique des langues est considérée comme un pilier majeur du TAL, en particulier pour la production de textes écrits. D'après Danlos et Roussarie (2000), il s'agit d'une branche du TAL qui vise à « produire des énoncés en langage naturel à partir de représentations informatisées » (p. 311). La génération est un processus inverse de l'analyse qui, lui, vise à extraire des textes leur « représentation » qui se définit comme « le résultat de l'analyse et le point de départ de la génération » qui rend « explicites différents types d'informations » de nature morphologique, syntaxique, sémantique ou pragmatique (Bouillon, 1998, p. 28).

5 Cette représentation peut être considérée comme un modèle de référence qui se dresse comme point de départ pour plusieurs opérations de traitement du langage naturel. On obtient une représentation par le biais d'un étiquetage qui infère les mécanismes linguistiques soit à partir de données brutes (étiquetage non supervisé) ou à

l'aide d'un jeu de données assorties d'une annotation (étiquetage supervisé). Comme l'expliquent Garside et coll. (1997), l'étiquetage consiste à ajouter des informations interprétatives de nature linguistique (par exemple d'ordre morphologique ou syntaxique) à un texte. Selon Bird et Liberman (2001), ces données peuvent être des informations « descriptives ou analytiques ». Dans la pratique, l'étiquetage s'effectue à l'aide d'outils adéquats – des étiqueteurs – comme, par exemple, TreeTagger (Schmid, 1997, 1999 ; Schmid & Laws, 2008) ou d'autres.

6 Dans la même lignée du développement du TAL, la démocratisation des outils informatiques, du web et des applications grand public accessibles via les ordinateurs personnels et les téléphones portables, est un processus parallèle et concomitant au développement de serveurs capables de stocker et de traiter de grandes masses de données (Big Data) dans un laps de temps relativement court (Sébillot, 2015). Dans l'analyse de Sébillot (2015), ces avancées technologiques ne constituent pas une « révolution », mais s'inscrivent dans une continuité historique (p. 43). Elles reposent sur la capacité des machines à traiter des quantités de plus en plus grandes de données qui peuvent être dégradées ou structurées. En revanche, ce qui apparaît comme inédit c'est qu'au cours des dix dernières années, nous assistons à une véritable révolution en termes d'accessibilité des données volumineuses, soit dans des ensembles structurés comme des corpus annotés, soit librement accessibles, accompagnés ou non d'outils pour le traitement et l'exploitation structurée. Dans cette perspective, l'ouverture au grand public des grands modèles de l'IA au début des années 2020 a donné lieu à de nombreuses initiatives pour repenser certains domaines de l'activité humaine, comme l'éducation, visant à former les enseignants et des concepteurs pédagogiques à leur utilisation.

1.2. L'intelligence artificielle en éducation : une affaire scientifique et politique au service des acteurs

7 L'usage des technologies de l'information et de la communication en éducation (TICE) est une problématique de recherche récurrente en didactique des langues. Les pratiques enseignantes qui s'en saisissent

et les représentations ont fait l'objet de nombreuses études dans le domaine (Burrows & Miras, 2019; Guichon, 2012 ; 2015). Elles ont permis, entre autres, de mettre en exergue les nombreuses interrogations que ces outils suscitent auprès de la communauté enseignante, allant de l'utilisation en classe à la constitution des corpus spécialisés ou généralistes au service de l'enseignement/apprentissage des langues et la conception des ressources qui en émanent (Cavalla, 2020; Ruggia & Gaillat, 2023), notamment en synergie avec les acquis du TAL (Antoniadis & coll., 2005). D'autres travaux de recherche ont donné lieu à la réalisation d'applications concrètes portant sur l'annotation des erreurs d'apprenants à partir de corpus (Granger, 2002, 2007). Par ailleurs, nous relevons de plus en plus de recherches autour de la création d'outils basés sur l'IA comme DeepFle (Ruggia, 2021).

8 Plus récemment, des études ont été menées pour identifier des pratiques et des représentations des outils de l'IA (Kaplan-Rakowski & coll., 2023 ; Son & coll., 2023). Le travail de Son et coll. (2023) fait mention de sept domaines dans lesquels l'IA peut trouver des applications concrètes :

- Le traitement automatique des langues
- L'apprentissage par corpus (« data driven learning »)
- L'évaluation automatique de l'écrit
- L'évaluation dynamique assistée par ordinateur
- Les systèmes de tutorat intelligents
- La reconnaissance automatique de la parole
- Les chatbots

9 La recherche dans le domaine de l'IAG a incité nombre de décideurs politiques et de chercheurs à se saisir de la question de l'usage des outils de l'IAG en éducation. Nous relevons, ainsi, le consortium AI4T (*Artificial intelligence for teachers* [5]) qui vise à identifier les représentations de l'IA en éducation et propose des pistes pour former la communauté enseignante à l'utilisation des outils issus de l'IA (Higuera & Iyer, 2024, p. 28-33). Cette initiative n'est qu'une des nombreuses réflexions menées dans la communauté scientifique en éducation à l'échelle européenne. Dans le contexte français, différents groupes de réflexion thématiques sur le numérique (GTN) du ministère de l'Éducation nationale structurent la réflexion autour de

travaux académiques mutualisés (TraAM [6]). Dans ces initiatives, nous avons repéré des travaux qui visent à optimiser la création des ressources pour l'enseignement des langues via Moodle comme, par exemple, dans les travaux de l'Académie de Strasbourg [7] dont la ligne directrice est de développer la littéracie numérique (Ollivier, 2018 ; Ollivier & coll., 2016), chez les enseignants et les élèves du secondaire. On y trouve des procédures de génération automatique d'activités pédagogiques comme, par exemple, des exercices grammaticaux à l'aide d'outils comme H5P [8] ou les modules natifs de Moodle comme le Test. Enfin, nous avons répertorié une initiative de génération automatique des contenus adaptés à Moodle de l'université de Friburg qui a organisé un webinaire sur ce sujet en 2024 [9] orienté vers la structure technique des outils et des supports produits.

1.3. Qu'en est-il de la conception pédagogique ?

10 La perspective de génération automatique des ressources pour les LMS fait écho à trois projets dans lesquels nous avons été étroitement impliqué : d'une part le projet PERL [10], créé dans la logique de la loi ESR 2013 visant à mettre en place une plateforme de ressources pour l'enseignement/apprentissage des langues pour les spécialistes d'autres disciplines (LANSAD) destinée aux dix universités de la COMUE Université Sorbonne Paris Cité (USPC) (Burrows & coll., 2019 ; Burrows & Miras, 2019), d'autre part le chantier Hybridation du programme PIA3 Nexus de l'Université de Montpellier Paul-Valéry qui vise la transformation des formations en LANSAD pour les 12 langues enseignées à l'Université (Tovar & coll., 2023) et, enfin, le projet Bolivia Habla Fr@ncés [11] (désormais BHF) mené conjointement par le réseau des Alliances françaises de la Bolivie et l'Université de Montpellier Paul-Valéry dont le but est de proposer une plateforme d'enseignement hybride du français dans une préparation du diplôme élémentaire de langue française (DELF). Au cours de ces expériences, nous avons eu l'occasion de mesurer la transformation du rôle de l'enseignant dont la posture se redéfinit et le pousse à opérer des choix du matériel pédagogique, voire de le concevoir.

11 Partant de ces expériences, nous nous inscrivons dans une démarche de recherche-conception/développement, au sens où la production

de ressources pédagogiques sert à la fois de support d'expérimentation et d'itération pour valider des hypothèses d'ordre didactique et ingénieurique. Comme le souligne Guichon (2007), la recherche-développement (R-D) en didactique des langues vise à dépasser le simple stade des préconisations théoriques. Elle consiste à concevoir, puis à tester de manière systématique un dispositif ou une ressource, avant de l'améliorer sur la base des retours utilisateurs. Cette logique hybride – entre production concrète (résultat tangible) et analyse scientifique (théorisation) – permet d'aboutir à des solutions pensées pour un contexte d'apprentissage précis (dans notre cas un environnement numérique de type Moodle), tout en générant de nouvelles connaissances au travers d'une démarche qui combine l'IAG et les outils du TAL. Elle repose sur un cycle itératif qui suppose un passage par la problématisation, la conception, l'implémentation et l'évaluation des ressources produites. Le protocole que nous détaillons s'inspire de la méthodologie décrite par Guichon : chaque étape nous permet à la fois d'affiner les résultats et de procéder à leur analyse, ce qui fait écho à la démarche ADDIE[12] (Gustafson & Branch, 2002 ; Musial & Tricot, 2020).

12 Ainsi, l'identité enseignante se voit-elle évoluer vers un rôle d'« enseignant-concepteur » (Burrows & Miras, 2019) qui n'a pas comme vocation unique de médiatiser le matériel pédagogique, mais de penser à toute une démarche d'ingénierie didactique qui vise à prendre en compte les contraintes selon lesquelles l'enseignement des langues se développe, notamment à l'université et dans les organismes de formation du réseau de coopération et d'action culturelle. Ce sont ces observations qui ont orienté en grande partie notre réflexion sur la capacité des outils de l'IAG, et notamment de ChatGPT, à automatiser la création des contenus adaptés à leurs besoins et optimisés pour l'enseignement des langues via Moodle. En d'autres termes, nous souhaitons voir si nous pouvons gagner du temps et « innover avec l'existant », pour reprendre l'expression de Sauvage (2019), plutôt que de nous lancer dans une multiplication des initiatives de conception pédagogique et, ainsi, proposer de nouvelles pistes pour une formation efficace des enseignants en langue.

2. Description des outils mobilisés et du protocole expérimental

13 Afin de garantir la pertinence didactique des ressources générées, nous avons défini un ensemble de critères articulant contraintes linguistiques, pédagogiques et techniques. Sur le plan linguistique, les productions ont été analysées à l'aune des critères de l'inventaire linguistique du CECRL (North & Fondation Eurocentres/Eaquals, 2015) qui donne des informations sur plusieurs caractéristiques du niveau visé comme, par exemple, les contenus lexicaux et grammaticaux, les tâches sociolangagières visées et les documents adaptés pour la mise en place d'activités pédagogiques. Le calibrage lexical a été vérifié à l'aide de FleLex et la génération d'activités de grammaire s'est effectué à partir de la version annotée à l'aide de TreeTagger. Sur le plan pédagogique, chaque ressource a été examinée selon sa fonction dans une unité didactique (document déclencheur, support lexical, tâche d'évaluation) et sa capacité à favoriser l'autonomie ou la progression. Enfin, la transférabilité vers Moodle via le format XML a également été intégrée à l'analyse.

14 Dans cette section, nous allons décrire les outils mobilisés et les étapes du protocole expérimental visant à tester notre hypothèse qui, pour rappel, stipule que la combinaison de ChatGPT, de FleLex et de TreeTagger permet de générer des ressources et des activités en accord avec le CECRL et compatible avec Moodle. Dans un premier temps, notre démarche prévoit de faire appel à une combinaison d'outils de supervision, tels que FleLex et TreeTagger, pour calibrer les ressources pédagogiques générées par ChatGPT en fonction des descripteurs du CECRL. Ensuite, nous détaillons l'intégration des ressources produites dans Moodle, via des activités spécifiques comme les tests et les glossaires, ce qui permet une intégration flexible grâce aux formats propices. L'évaluation des ressources produites est faite, donc, à partir de deux critères principaux : leur adéquation aux contenus linguistiques du CECRL et leur compatibilité avec le LMS Moodle.

2.1. Les outils de supervision : le CECRL, FleLex et TreeTagger

15 Comme nous l'avons vu dans la section 1.1, l'étiquetage est une opération qui peut être supervisée ou non supervisée. Cette même distinction peut être appliquée au protocole que nous testons : nous qualifierons la génération supervisée de ressources comme une production qui fait appel à des ressources (North & Fondation Eurocentres/Eaquals, 2015 ; *Conseil de l'Europe*, 2001) dans le but de calibrer les documents produits au niveau visé. À l'inverse, lorsque la génération s'effectue sans faire appel à ces ressources, nous sommes face à une génération non supervisée. Dans cette partie, nous allons faire un panorama sur les outils que nous utilisons dans la génération supervisée de ressources pour les besoins du protocole que nous décrirons dans la section 3.4.

16 Document de référence de la perspective actionnelle (PA), le CECRL structure l'apprentissage d'une langue en six niveaux de référence qui correspondent à la capacité d'un apprenant à réaliser des tâches communicatives (Conseil de l'Europe, 2001, p. 32) ancrées dans quatre grands domaines : public, professionnel, éducationnel et personnel (p. 18). Il s'agit d'un outil complexe qui permet de mettre en place une approche hiérarchisée à travers des descripteurs dont la création dépend de plusieurs critères. L'utilité des descripteurs va au-delà de l'apprentissage, de l'enseignement et de l'évaluation et permet, également, de concevoir le matériel pédagogique pour calibrer les programmes qui visent l'obtention d'un niveau en langue. Dans cette perspective, nous avons étudié les programmes du DELF et du DALF en lien avec les descripteurs du CECRL qui ont été traduits dans un inventaire linguistique (North & Fondation Eurocentres/Eaquals, 2015). Ce document propose des scénarios types, permettant de prévoir les points à prendre en compte pour développer la compétence à communiquer langagièrement, mais aussi des informations sur les types de textes qui pourraient être employés en tant que documents déclencheurs au sein d'une unité didactique d'enseignement/apprentissage du FLE (North & Fondation Eurocentres/Eaquals, 2015, p. 42), des éléments de grammaire, ainsi que les principales tâches sociolangagières (fonctions). Les critères de cet

inventaire ont été, en effet, le premier élément que nous avons choisi pour générer les ressources produites.

17 Le deuxième outil pour la supervision – FleLex (Francois & coll., 2014 ; Pintard & François, 2020) est un outil qui permet d'effectuer une analyse lexicale des textes au regard des niveaux visés du CECRL et de nous fournir des informations sur le niveau auquel un texte est adapté. Cette analyse s'effectue à partir d'une lemmatisation suivie d'une comparaison de la distribution du lexique par niveau à partir d'une liste de fréquence. Cette liste a été établie à partir de l'étude de 2071 textes issus de 28 manuels de FLE représentant un corpus de près de 777 000 mots. Le corpus de référence est annoté à l'aide de TreeTagger et il contient des informations sur près de 14 500 mots qui servent de repère pour une classification lexicale. L'outil est doté d'une interface web et son utilisation est gratuitement offerte à la communauté scientifique et aux enseignants de FLE. Les utilisateurs peuvent soit rechercher la fréquence d'une unité lexicale par niveau, soit analyser la structure lexicale d'un texte en vue de l'utiliser dans une perspective d'enseignement du FLE. Dans le cadre de notre protocole, nous utilisons FleLex pour la partie supervision, qui nous permet d'analyser et de calibrer le lexique des ressources générées par ChatGPT.

18 Enfin, TreeTagger (Schmid, 1997, 1999) est un étiqueteur morphosyntaxique couramment utilisé en TAL et en linguistique de corpus. Il assigne une classe grammaticale à partir d'un set d'étiquettes (*tagset*) et un lemme (forme lexicale) à chaque token (mot graphique), selon des modèles statistiques préalablement entraînés pour 35 langues. Pour le français, ces étiquettes concernent les différentes classes de mots, mais aussi la ponctuation (qui inclut aussi des citations), aussi bien au milieu qu'à la fin des phrases, des symboles, des pronoms, des noms propres et des numéraux (ordinaux et cardinaux). Une version web de TreeTagger permet d'analyser rapidement des petits échantillons de textes en ligne ou de téléverser des fichiers et d'obtenir des résultats identiques sous forme d'un fichier que l'utilisateur peut exploiter ultérieurement. Dans notre protocole, nous utilisons TreeTagger spécifiquement dans le but de faciliter la génération d'activités relatives à la grammaire et à la compréhension écrite : chaque texte est étiqueté par l'outil en ligne et c'est le résultat obtenu par

TreeTagger qui est convoqué pour générer les ressources et les activités.

2.2. Formats des ressources et des activités générées

19 Les ressources et les activités pédagogiques générées dans notre expérimentation prennent différents formats : d'une part, nous avons des ressources générées au format texte et image et, d'autre part, des ressources adaptables au LMS Moodle sous forme d'activités (tests) ou de ressources lexicales (glossaires). En effet, les outils « Test » et « Glossaire » se distinguent par leur capacité à structurer et organiser des contenus pédagogiques à l'aide de formats très flexibles comme le format XML (Michard, 2000). Ce format, adapté à la manipulation de chaînes de caractères, permet de structurer des contenus complexes tout en facilitant leur réutilisation ou leur modification et le transfert vers d'autres plateformes du même type, ce qui garantit l'interopérabilité des ressources générées. Notre choix de nous focaliser principalement sur les activités « Test » et « Glossaire » est motivé par la génération assez aisée du langage XML à l'aide de ChatGPT : en effet, la création manuelle sur une plateforme Moodle de ce type d'activités et de ressource est très chronophage.

20 La structure d'un fichier XML de test Moodle repose sur des balises clés comme <quiz> et <question>, permettant de définir des catégories, des types de questions (choix multiple, vrai/faux, réponse courte, etc.), et leurs formats spécifiques. L'utilisation du langage XML simplifie également la gestion des activités pédagogiques complexes grâce à des balises pour les réponses attendues et les rétroactions. Le glossaire est un autre outil qui permet de tirer parti des avantages du format XML pour transférer et gérer des articles que l'on peut mobiliser dans des activités qui demandent une accessibilité au lexique. La structure XML du glossaire est relativement simple : pour chaque entrée (<entry>) le glossaire dispose d'un terme (<concept>) pour lequel il faut insérer une définition (<definition>) qui peut également contenir d'autres informations comme des traductions ou des informations sur la classe des mots et, dans certains cas, leur prononciation. Lors de l'exportation, un fichier structuré contenant les entrées du glossaire est généré, tandis que l'importation

offre une grande flexibilité pour ajouter des contenus dans un glossaire existant ou en créer un nouveau. Bien que ce processus semble relativement aisé, il requiert une attention particulière à la structure du fichier XML, notamment pour éviter des erreurs liées aux balises mal formatées ou aux doublons. Il peut être vérifié au travers des outils de diagnostic XML en ligne comme, par exemple, XML Validator[13].

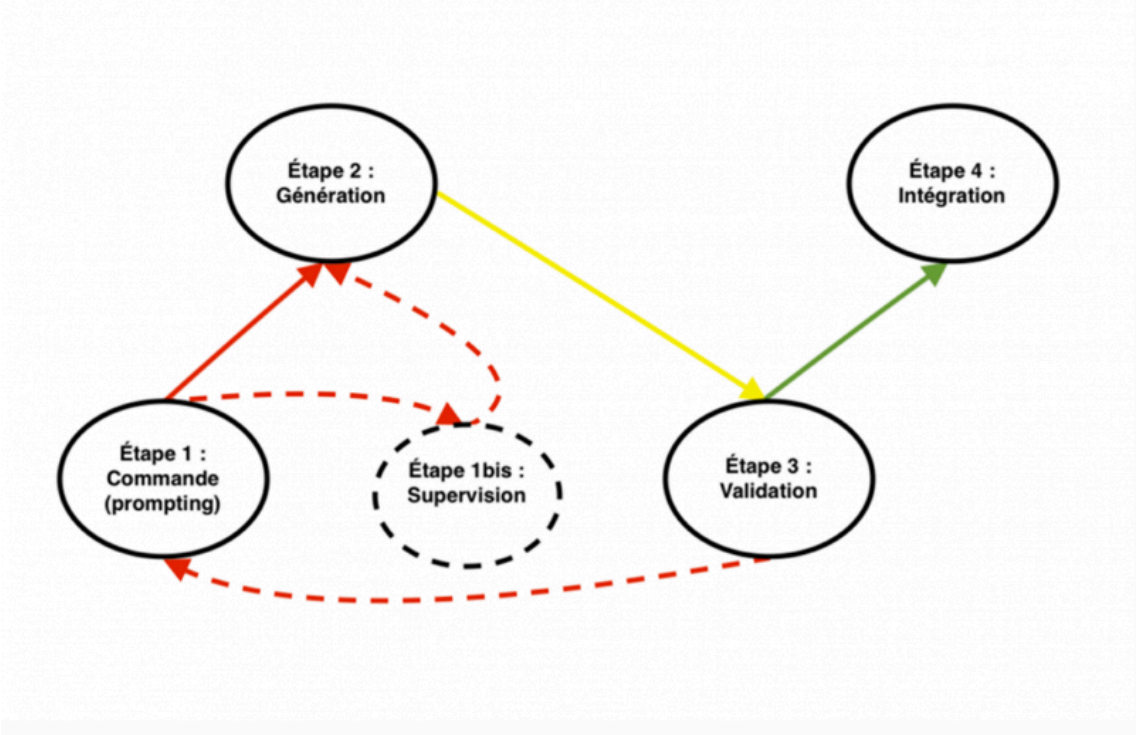
2.3. Génération des ressources textuelles et iconographiques

21 Dans le cadre d'une unité didactique, les documents (authentiques ou construits) sont principalement utilisés comme déclencheurs d'un cycle d'apprentissage. De surcroît, la première étape de notre protocole de création de ressources avec ChatGPT est de demander à l'outil de générer des documents calibrés pour les niveaux A1 et A2 de manière non supervisée, avec une ligne de commande[14] explicite et concise, puis de vérifier la conformité de ces documents avec le niveau visé sur le plan linguistique, à l'aide de l'inventaire linguistique des descripteurs du CECRL et de FleLex.

2.4. Génération et adaptation des activités pédagogiques sur Moodle

22 Cette seconde étape vise à créer des activités pédagogiques axées sur la compréhension et la production écrite, sur la grammaire et sur le lexique. Le point de départ de cette étape est constitué des ressources textuelles et iconographiques créées dans l'étape précédente. À cette étape, les activités sont produites soit à partir de documents bruts, soit à partir de documents annotés avec TreeTagger (notamment pour les activités relatives à la grammaire et des glossaires), et testées dans différentes plateformes Moodle afin de vérifier la compatibilité avec les formats. Notre protocole peut être schématisé de la façon suivante :

Figure 1 : Étapes du protocole de création des ressources et d'activités



23 Les lignes de commande qui nous ont permis de générer les ressources figurent dans les exemples que nous analysons dans la section suivante.

3. Test du protocole et analyse des résultats

24 La mise en œuvre du protocole expérimental nous a permis de générer une série de ressources pédagogiques pour les niveaux A1 et A2 du CECRL. Dans cette section, nous allons analyser ces productions sous deux angles principaux : d'une part, la qualité et la variété des ressources générées par les outils d'IAG au regard de leur contenu et du niveau visé et, de l'autre, leur compatibilité avec Moodle. Ces analyses nous permettront d'évaluer la viabilité et les limites du protocole proposé, tout en identifiant des pistes pour son amélioration future.

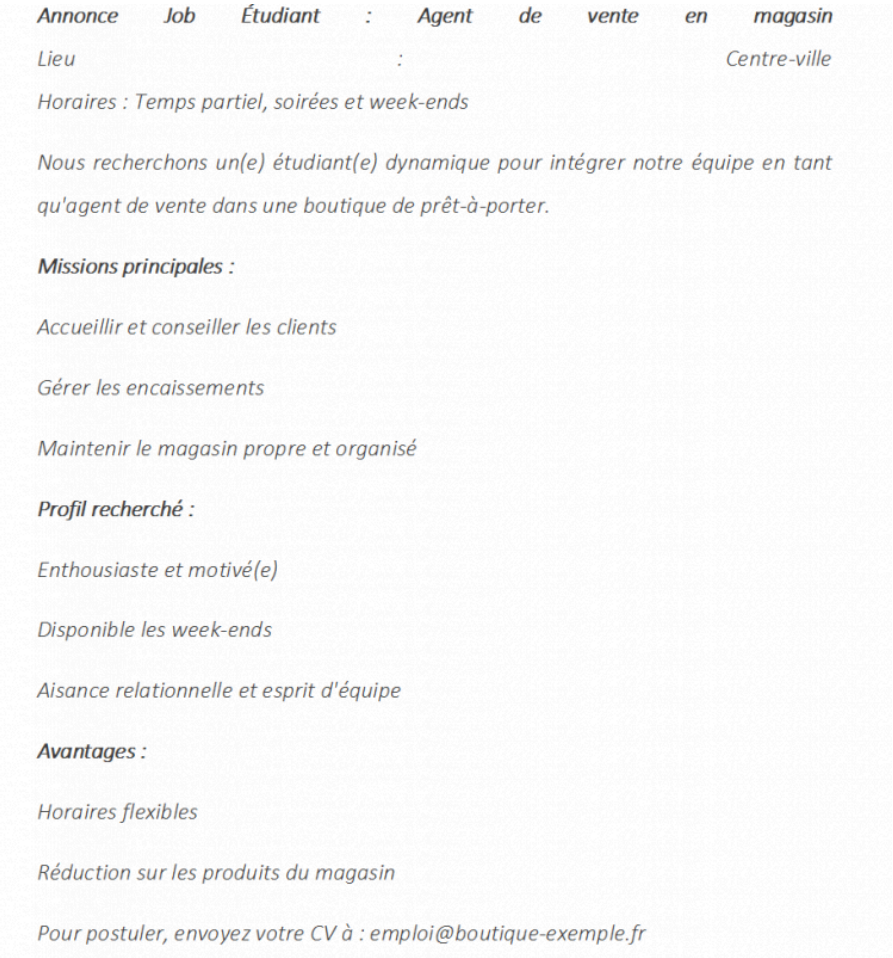
3.1. Génération des ressources textuelles et iconographiques

25 Parmi les types de documents pour le niveau A1 qui figurent dans l'inventaire linguistique du CECRL, on trouve l'annonce qui représente un document déclencheur assez pertinent pour amener les apprenants à construire une présentation (qui peut, par exemple, prendre la forme d'une lettre de motivation ou d'un *curriculum vitæ*) ou, le cas échéant, un courriel envoyé à l'employeur pour demander des précisions et renseignements complémentaires sur l'emploi visé. Nous avons demandé à ChatGPT de nous construire une annonce pour un job étudiant avec la ligne de commande suivante :

> Peux-tu me générer un texte qui ressemble à une annonce de travail ? Cela peut être des annonces simples pour un job étudiant, par exemple.

26 Le résultat obtenu semble convenable au regard de l'inventaire du CECRL. Il se présente sous la forme suivante :

Figure 2 : Génération d'un document déclencheur textuel pour le niveau A1



Annonce Job Étudiant : Agent de vente en magasin

Lieu : Centre-ville

Horaires : Temps partiel, soirées et week-ends

Nous recherchons un(e) étudiant(e) dynamique pour intégrer notre équipe en tant qu'agent de vente dans une boutique de prêt-à-porter.

Missions principales :

Accueillir et conseiller les clients

Gérer les encaissements

Maintenir le magasin propre et organisé

Profil recherché :

Enthousiaste et motivé(e)

Disponible les week-ends

Aisance relationnelle et esprit d'équipe

Avantages :

Horaires flexibles

Réduction sur les produits du magasin

Pour postuler, envoyez votre CV à : emploi@boutique-exemple.fr

27 Nous avons également demandé à ChatGPT de nous générer une image qui illustre cette annonce avec la ligne de commande suivante :

> Peux-tu m'illustrer le texte avec une image ?.

28 Le résultat obtenu est le suivant :

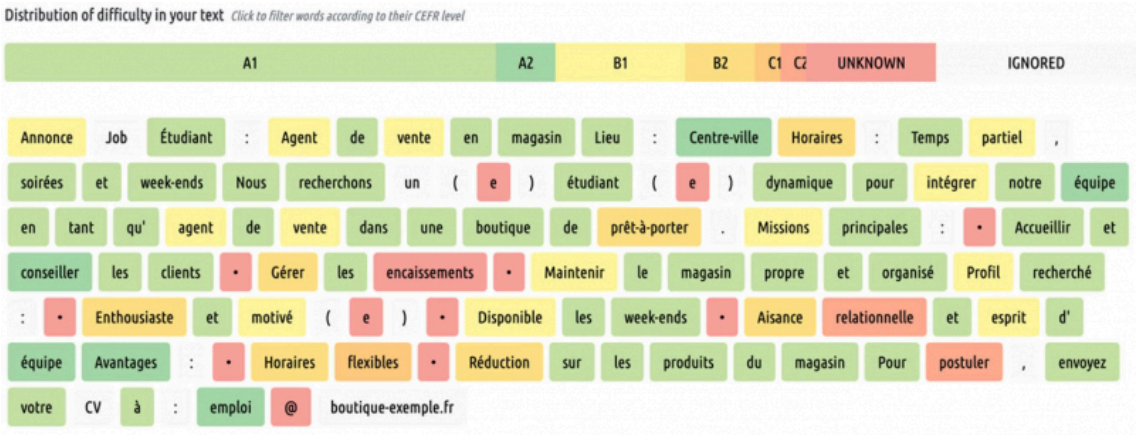
Figure 3 : Génération d'une image en lien avec le document déclencheur textuel A1



29 A priori, l'image générée par ChatGPT illustre bien le texte de l'annonce et peut, en effet, renforcer le document déclencheur, en jouant un rôle à la fois illustratif et heuristique. Sur le plan cognitif, elle favorise l'activation des connaissances antérieures, l'anticipation lexicale et la projection dans la situation communicative. Dans le cadre de tâches actionnelles, elle permet d'initier des dialogues, des descriptions ou des prises de position, en lien avec des actes de parole ciblés par le CECRL. Cette complémentarité texte/image est également cohérente avec le principe multimodalité (Mayer, 2017, 2020), favorisant ainsi la compréhension et la mémorisation.

30 Lorsque nous analysons le texte à l'aide de FleLex d'un point de vue du lexique, nous observons le résultat suivant :

Figure 4 : Mesure de la complexité lexicale du texte produit à l'aide de FleLex.



31 Selon l'outil, la distribution des mots par niveau du CECRL est la suivante :

Tableau 1 : Distribution des mots de l'exemple généré pour le niveau A1

Niveau	Nombre de mots	Pourcentage
A1	49	44,14 %
A2	6	5,41 %
B1	13	11,71 %
B2	7	6,31 %
C1	1	0,90 %
C2	2	1,80 %
Niveau non défini	13	11,71 %
Mots ignorés	20	18,20 %

32 Nous avons procédé à quelques ajustements (ex. conseiller → donner des conseils ; gérer les encaissements → gérer la caisse) ce qui nous a permis d'obtenir un score de mots adaptés au niveau A1 supérieur à celui initialement relevé (52,3 %). Ce score s'explique notamment par le caractère très particulier de ce type de texte : son architecture privilégie des informations sous forme d'une liste à puces. Nous avons testé d'autres types de textes qui présentent des scores supérieurs : environ 50 % de mots de niveau A1 pour une recette de cuisine simple, pour un menu ou pour un prospectus simple. Néanmoins, cela

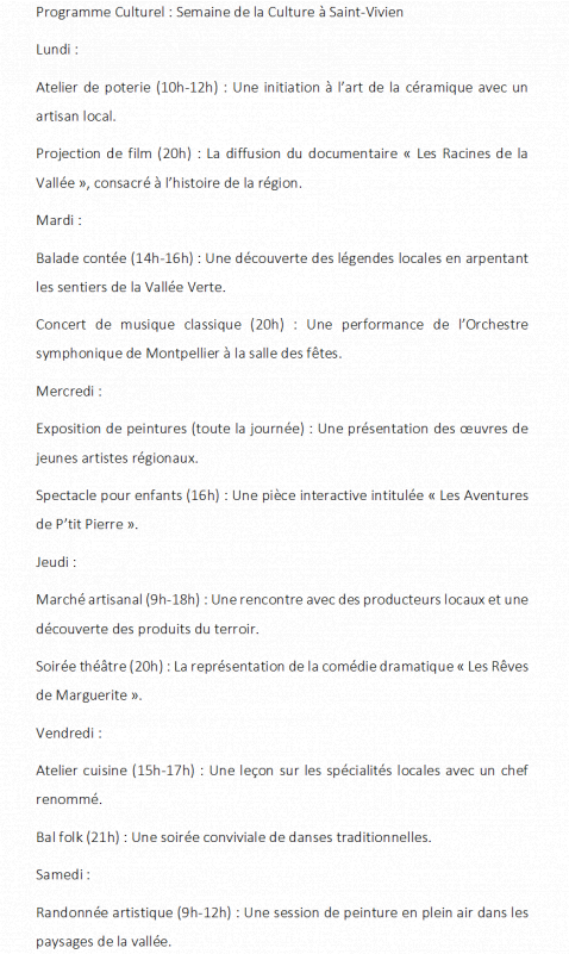
ne nous paraît pas problématique et la présence de mots habituellement abordés dans des niveaux supérieurs (A2 ou B1) peut se justifier par l'intention d'amener les apprenants à s'approprier progressivement le lexique spécifique à la thématique abordée, en évitant, toutefois, d'aller vers les mots de niveau B2 et C1, sans pour autant les restreindre de manière catégorique. Le meilleur score est obtenu lors de la génération des récits courts (66,49 % des mots) et des anecdotes (64,71 % des mots).

33 Concernant le niveau A2, nous avons demandé la génération d'un document de type programme d'événements culturels avec la ligne de commande suivante :

> Peux-tu me proposer un programme d'événements culturels ?

34 Le résultat obtenu est le suivant :

Figure 5 : Génération d'un document déclencheur textuel pour le niveau A2



Programme Culturel : Semaine de la Culture à Saint-Vivien

Lundi :

Atelier de poterie (10h-12h) : Une initiation à l'art de la céramique avec un artisan local.

Projection de film (20h) : La diffusion du documentaire « Les Racines de la Vallée », consacré à l'histoire de la région.

Mardi :

Balade contée (14h-16h) : Une découverte des légendes locales en arpentant les sentiers de la Vallée Verte.

Concert de musique classique (20h) : Une performance de l'Orchestre symphonique de Montpellier à la salle des fêtes.

Mercredi :

Exposition de peintures (toute la journée) : Une présentation des œuvres de jeunes artistes régionaux.

Spectacle pour enfants (16h) : Une pièce interactive intitulée « Les Aventures de P'tit Pierre ».

Jeudi :

Marché artisanal (9h-18h) : Une rencontre avec des producteurs locaux et une découverte des produits du terroir.

Soirée théâtre (20h) : La représentation de la comédie dramatique « Les Rêves de Marguerite ».

Vendredi :

Atelier cuisine (15h-17h) : Une leçon sur les spécialités locales avec un chef renommé.

Bal folk (21h) : Une soirée conviviale de danses traditionnelles.

Samedi :

Randonnée artistique (9h-12h) : Une session de peinture en plein air dans les paysages de la vallée.

35 Dans la même logique que pour le niveau A1, nous avons également demandé à ChatGPT de nous générer une image illustrant ce texte avec la ligne de commande suivante :

> Peux-tu m'illustrer ce texte avec une image ?

36 Le résultat obtenu est le suivant :

Figure 6 : Image générée par ChatGPT en lien avec le document A2



37 Après analyse de ce texte avec FleLex, nous obtenons les résultats suivants :

Tableau 2 : Distribution des mots de l'exemple généré pour le niveau A2

Niveau	Nombre de mots	Pourcentage
A1	161	53,4 %

A2	7	2,27 %
B1	16	5,18 %
B2	14	4,53 %
C1	5	1,62 %
C2	3	0,97 %
Niveau non défini	29	9,39 %
Mots ignorés	74	23,95 %

38 Concernant le lexique dans le document généré, nous constatons une majorité de mots qui correspondent davantage au niveau inférieur, tandis que les mots qui correspondent, selon FleLex, au niveau A2 sont très peu nombreux et même inférieur à ceux des niveaux B. Rappelons que FleLex est un outil qui ne mesure pas, mais qui donne des indications sur la fréquence des mots par niveau en fonction d'un matériel d'entraînement : les mots et leurs formes peuvent apparaître à n'importe quel niveau de langue. C'est pour cela qu'il est plus pertinent de considérer que ce texte se compose de mots qui sont très fréquents au niveau A1 et qu'il introduit progressivement des mots de niveau A2, B1 et B2. Pour les autres types de documents, nous avons obtenu des scores similaires : pour le récit qui se veut un texte un peu plus élaboré l'outil affiche 61 % de mots de niveau A1 et 5,47 % de mots de niveau A2, tandis que pour l'anecdote qui est similaire au récit, nous obtenons 64 % de mots de niveau A1 et 4,66 % de mots fréquents au niveau A2.

39 Quant à l'image générée, il est à noter qu'il s'agit d'un tableau qui peut potentiellement être utilisé mais qui contient des mots orthographiquement proches de ce qui est abordé dans le texte : ceux-ci peuvent prêter à confusion (ex. « artison » pour « artisan »). Pour cette raison, nous avons spécifié à l'outil de nous générer une image qui ne contient pas de texte avec la ligne de commande suivante :

> Peux-tu me régénérer une image qui illustre ce texte sans utiliser des mots ni des lettres ?

40 Le résultat obtenu est, alors, le suivant :

Figure 7 : Image générée par ChatGPT en lien avec le document A2 (ajustement)



41 En effet, la spécification de la ligne de commande a permis de générer une image conforme à la demande qui illustre les différentes notions et événements abordés dans le texte (ateliers, bals, spectacles de théâtre, marché artisanal, etc.) sans éléments textuels qui pourraient potentiellement induire les apprenants en erreur.

3.2. Génération et analyse des activités de compréhension, de grammaire et des ressources lexicales

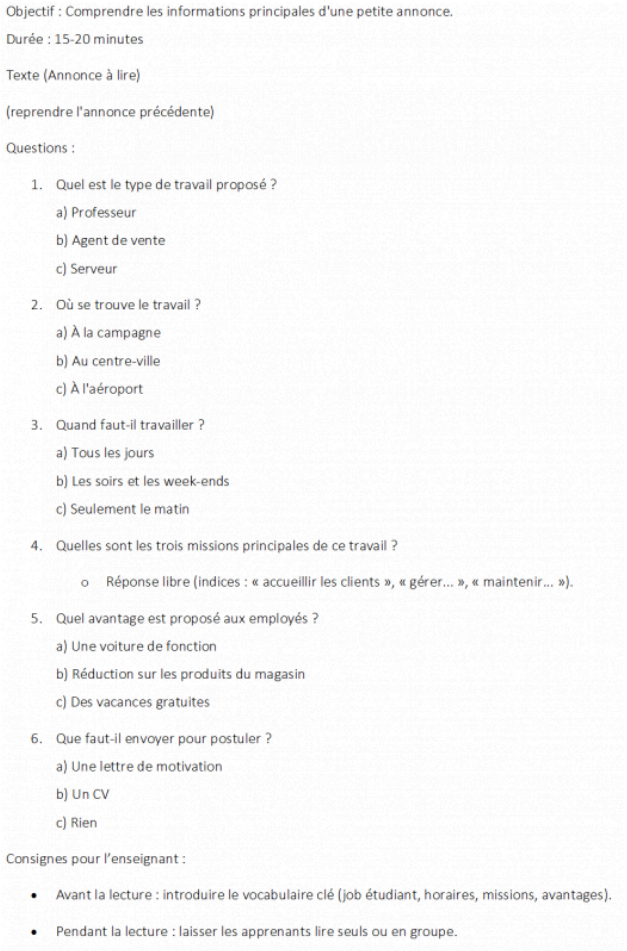
3.2.1. Activités de compréhension écrite

42 En lien avec les documents générés, nous avons demandé la création d'une batterie d'activités de compréhension écrite conformes aux

exigences du DELF A1. Ces activités sont majoritairement au format QCU/QCM [15]. Voici la ligne de commande utilisée et le résultat obtenu :

> Maintenant, est-ce que tu peux me proposer une activité de compréhension écrite de cette annonce pour un apprenant de niveau A1 en FLE ?

Figure 8 : Activité de compréhension écrite du document généré et proposition d'un scénario pédagogique



Objectif : Comprendre les informations principales d'une petite annonce.

Durée : 15-20 minutes

Texte (Annonce à lire)

(reprendre l'annonce précédente)

Questions :

1. Quel est le type de travail proposé ?
   a) Professeur
   b) Agent de vente
   c) Serveur
2. Où se trouve le travail ?
   a) À la campagne
   b) Au centre-ville
   c) À l'aéroport
3. Quand faut-il travailler ?
   a) Tous les jours
   b) Les soirs et les week-ends
   c) Seulement le matin
4. Quelles sont les trois missions principales de ce travail ?
   - Réponse libre (indices : « accueillir les clients », « gérer… », « maintenir… »).
5. Quel avantage est proposé aux employés ?
   a) Une voiture de fonction
   b) Réduction sur les produits du magasin
   c) Des vacances gratuites
6. Que faut-il envoyer pour postuler ?
   a) Une lettre de motivation
   b) Un CV
   c) Rien

Consignes pour l'enseignant :

- Avant la lecture : introduire le vocabulaire clé (job étudiant, horaires, missions, avantages).
- Pendant la lecture : laisser les apprenants lire seuls ou en groupe.

Pour ces résultats, notre constat est positif : en plus de l'activité qui est entièrement compatible avec Moodle, l'outil nous a également fourni un scénario qui permet de travailler à partir des documents déclencheurs. En dehors de la compréhension, ces activités abordent également le lexique en convoquant des mots initialement appris,

mais s'appuient aussi sur des outils spécifiques de Moodle sur lesquels nous allons revenir dans la suite de cet article.

3.2.2. Activités de grammaire

44 Nous avons également demandé à ChatGPT de nous générer des exercices grammaticaux en spécifiant les points de grammaire du niveau A1 et A2 à partir de l'inventaire du CECRL, et plus particulièrement de l'annexe E (North & Fondation Eurocentres/Eaquals, 2015, p. 56-58). L'idée principale de cette étape était de voir si l'outil était capable de nous générer des activités permettant de conceptualiser et de systématiser les faits de langue visés (« fonctions ») dans les deux niveaux. La génération s'est effectuée tout d'abord avec pour seule spécification le tableau des contenus linguistiques des niveaux A1 et A2 et le résultat semble pauvre d'un point de vue de la typologie des exercices avec, parfois, uniquement des vrai/faux, des QCU/QCM, des réponses courtes et des appariements. Pour obtenir un résultat satisfaisant et davantage axé sur les faits de langue observés au niveau visé, nous nous sommes tout d'abord orientés vers TreeTagger afin d'effectuer un étiquetage morphosyntaxique des textes produits et de mieux cibler les exercices :

Figure 9 : Extrait du texte de la petite annonce étiqueté à l'aide de TreeTagger

```
Nous    PRO:PER         nous
recherchons VER::pre rechercher
un      NUM     un
(       PUN     (
e       NOM     e
)       PUN     )
étudiant VER::ppr étudier
(       PUN     (
e       VER:pper        e
)       PUN     )
dynamique       ADJ     dynamique
pour    PRP     pour
intégrer VER::inf intégrer
notre   DET:POS         notre
équipe  NOM     équipe
en      PRP     en
tant    ADV     tant
qu' KON que
agent   NOM     agent
de      PRP     de
vente   NOM     vente
dans    PRP     dans
une     DET:ART         un
boutique        NOM     boutique
de      PRP     de
prêt-à-porter   NOM     prêt-à-porter
.       SENT    .
Missions        NOM     mission
principales     ADJ     principal
:       PUN     :
•       PUN     •
Accueillir VER::inf accueillir
et      KON     et
conseiller VER::inf conseiller
les     DET:ART         le
clients NOM     client
•       PUN     •
```

45 Bien que TreeTagger offre un étiquetage robuste et rapide, des erreurs ponctuelles ont été observées, notamment sur des mots polysémiques ou hors contexte ou, encore, sur les mots en écriture inclusive. Par exemple, le mot étudiant a parfois été analysé comme participe présent plutôt que comme nom, tandis que la marque du féminin dans un nom à écriture inclusive est traitée comme un verbe. Ce type d'ambiguïté nécessite une relecture humaine systématique, notamment lorsque l'étiquetage sert de base pour générer des exercices grammaticaux, dans le but de garantir la fiabilité des ressources proposées. Pour arriver à un résultat satisfaisant lors de la génération des exercices, nous avons spécifié qu'ils devaient concerner avant tout les tâches sociolangagières de chaque niveau visé en lien avec l'inventaire du CECRL (North & Fondation Eurocentres/Eaquals, 2015, p. 44) et que le nombre d'items devait être au minimum 4 pour

assurer un certain équilibre. La ligne de commande que nous avons utilisée est la suivante :

> Voici un fichier avec un texte étiqueté au format TreeTagger. Peux-tu me générer des exercices concernant les classes ouvertes de mots (nom, adjectif, verbe, adverbe) en lien avec ce texte ? Pour chaque exercice, il faut 4 items minimum et on peut limiter la création à 10 exercices au total.

ChatGPT nous a fourni une batterie de dix exercices permettant à l'apprenant d'identifier une classe de mots, de transformer des mots du masculin vers le féminin et *vice-versa*, de compléter des phrases avec des mots manquants (notamment des verbes qui sont donnés à l'infinitif et que l'apprenant doit conjuguer lors de la complétion) ou, encore, d'identifier la classe des mots. La variété des exercices est satisfaisante et l'outil respecte les spécifications apportées comme en témoignent les exemples suivants (extrait) :

Figure 10 : Activité de compréhension écrite du document généré et proposition d'un scénario pédagogique

Exercice 2 – Mets ces adjectifs au féminin

Change ces mots pour les mettre au féminin :

Enthousiaste ? _____ (inchangé)

Partiel ? _____

Principal ? _____

Relationnel ? _____

Exercice 6 – Complète ces phrases avec l'adjectif correct

Écris correctement les mots entre parenthèses :

Les horaires sont très _____. (flexible)

Une personne _____. (enthousiaste)

Voici les missions _____. (principal)

Une aisance _____. (relationnel)

47 Nous avons relevé quelques problèmes potentiels comme, par exemple, les consignes qui peuvent paraître trop techniques pour certains apprenants de FLE ne maîtrisant pas forcément le métalangage en français, ni les verbes de ce que l'on appelle les « discours procéduraux » (Garcia-Debanc, 2001 ; Pery-Woodley, 2001) dont fait partie la consigne. Pour cela, nous avons demandé à l'outil d'utiliser les termes grammaticaux avec parcimonie. Une fois cette étape validée, nous avons demandé à l'outil de nous générer une batterie de questions compatible avec Moodle de sorte à pouvoir l'injecter dans une banque de questions que nous avons testée sur la plateforme BHF. Le résultat est relativement satisfaisant, mais quelques points de vigilance sont à observer dans le but de parfaire le rendu.

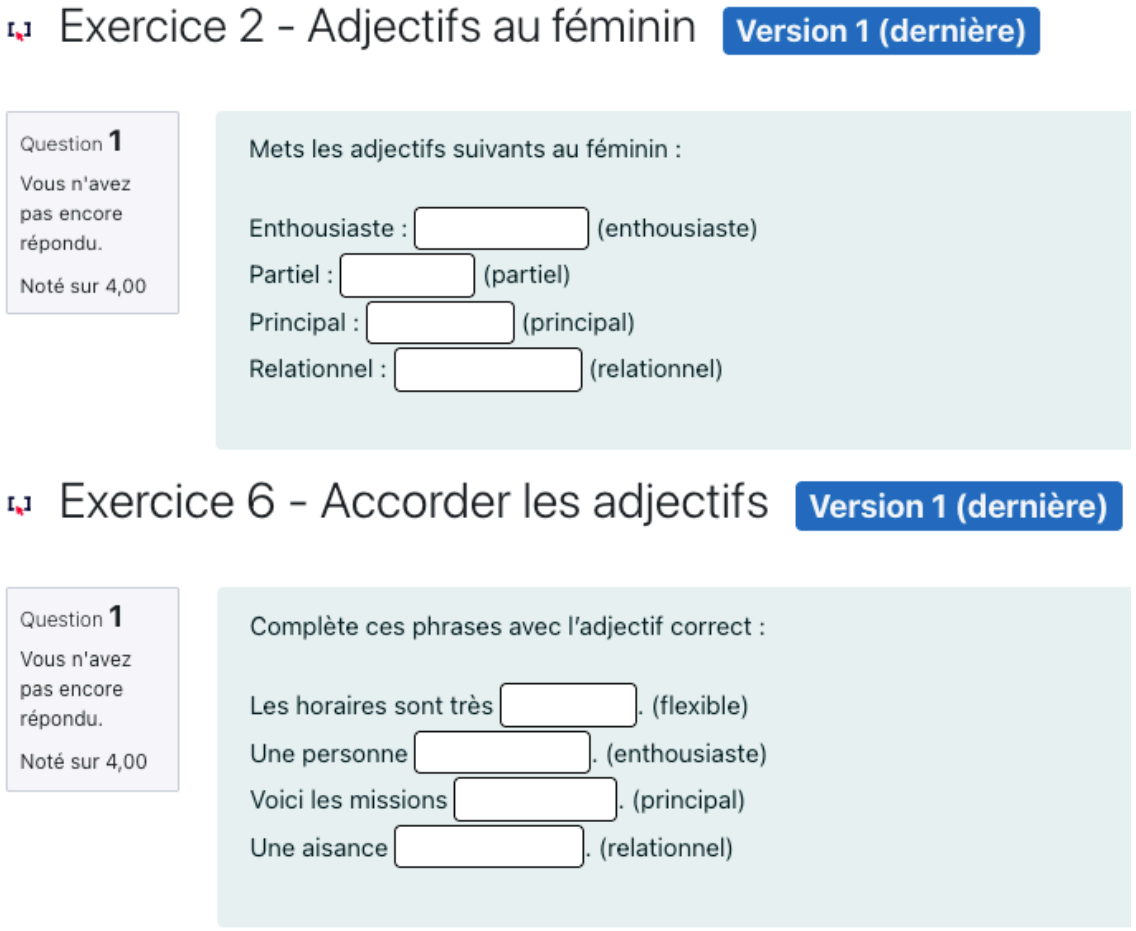
48 Ce que nous relevons comme une difficulté majeure dans la conception de ces activités est leur transformation au format XML compatible avec Moodle. En effet, nous avons constaté que ChatGPT fournit, dans certains cas, des modèles de transformation et des exemples que l'utilisateur doit suivre, mais refuse, même avec la version payante, de générer des fichiers XML très longs, à l'exception du modèle 4o qui lui effectue un raisonnement « profond » et génère la totalité des exercices. Dans le cas de la version gratuite, qui a été utilisée pour cette expérimentation, l'utilisateur est invité à compléter les scripts XML lui-même, ce qui est très chronophage et, dans une certaine mesure, complexe, notamment pour ceux qui ne sont pas familiers avec les formats XML des différents types d'exercices de l'activité test qui peuvent varier d'une plateforme à une autre, suivant la configuration technique[16].

49 Des erreurs de formatage de la structure du XML pour chaque type d'exercices peuvent également survenir, ce qui empêche l'intégration des activités générées dans Moodle. Pour cela, il nous semble nécessaire d'effectuer le travail en deux temps, en spécifiant à l'outil les paramètres des différents exercices de la typologie de Moodle (étape d'entraînement) en lui fournissant un exemple pour chaque type d'exercice qui contient également les paramétrages souhaités. Pour faciliter cela, il est tout à fait possible de concevoir un exercice avec les paramétrages adéquats et de le télécharger depuis la banque de questions de Moodle au format XML : l'utilisateur peut, alors, copier la structure du fichier XML et demander à l'outil de reproduire le même paramétrage, comme dans l'exemple suivant :

> Peux-tu me reproduire l'exercice 2 et l'exercice 6 au format GAPFILL de Moodle ? Il faut suivre l'exemple suivant : [paramétrage d'un exercice de type au format XML]

50 En entraînant l'outil avec des modèles fournis, nous obtenons des résultats améliorés, ce dont témoignent les exemples suivants :

Figure 11 : Interface graphique de deux exercices de type *Gapfill* générés à l'aide de ChatGPT



3.2.3. Ressources lexicales

51 Un autre type de ressources qui peut être généré de façon optimale est le module Glossaire de Moodle. Ce module permet de constituer une liste de termes et de concepts, puis de les « appeler » (via des filtres) dans différentes activités pour offrir à l'apprenant un accès immédiat au sens des mots qu'il rencontre. Cet accès peut se faire depuis des ressources statiques (pages, leçons) ou dans des activités dynamiques (tests, forums), sans quitter la page consultée. Des glossaires plus élaborés peuvent aussi intégrer du son ou de l'image, afin d'inclure la prononciation ou d'illustrer visuellement un concept. Il est à noter que le glossaire propose un champ « termes liés », que l'on peut utiliser pour recenser, par exemple, des synonymes ou lister les

formes fléchies d'un mot, ce qui facilite l'accès au sens dans différents contextes.

52 Pour tester la génération automatique de glossaires, nous avons réutilisé les textes étiquetés à l'aide de TreeTagger afin d'établir une liste des mots et une traduction en espagnol à l'aide de la commande suivante :

> Peux-tu extraire la liste complète des mots de ce texte et me proposer une traduction en espagnol pour chacun d'entre eux ? Tu peux ranger le résultat au format suivant : « mot en français » → « traduction du mot en espagnol ».

53 ChatGPT génère, effectivement, un résultat 100 % conforme à la demande en traduisant l'ensemble des mots du texte en espagnol. À partir de ce résultat, nous avons demandé à l'outil de nous construire un glossaire français – espagnol à l'aide de la ligne de commande suivante :

> Peux-tu générer un glossaire XML compatible pour Moodle pour l'ensemble des mots traduits en respectant la structure suivante : [échantillon d'un glossaire au format XML].

54 Le résultat obtenu est le suivant (reproduction partielle des premières entrées) :

Figure 12 : Glossaire Moodle des termes de l'annonce (A1) généré à l'aide de ChatGPT

```
<?xml version="1.0" encoding="UTF-8"?>
<GLOSSARY>
[...]
<ENTRIES>
<ENTRY><CONCEPT>annonce</CONCEPT><DEFINITION>anuncio</DEFINITION><FORMAT>1</FOR
MAT></ENTRY>
<ENTRY><CONCEPT>job</CONCEPT><DEFINITION>trabajo</DEFINITION><FORMAT>1</FORMAT><
/ENTRY>
<ENTRY><CONCEPT>étudiant</CONCEPT><DEFINITION>estudiante</DEFINITION><FORMAT>1</FO
RMAT></ENTRY>
[...]
```

55 Le glossaire obtenu a été injecté sur la plateforme BHF : en activant le filtre de liaison automatique [17], l'accès au sens des mots est immédiat depuis n'importe quelle activité. Ainsi, l'interprétation du lexique devient automatiquement accessible et l'apprenant peut consulter la définition d'un terme en un simple clic, comme l'illustre la figure 13 ci-après :

Figure 13 : Consultation du sens d'un mot dans un glossaire à partir d'une activité de compréhension écrite (testé sur la plateforme Bolivia Habla Fr@ncés).



56 Dans l'exemple présenté issu d'une activité de compréhension écrite où l'apprenant est amené à lire le texte et à répondre à une série de questions, la lecture est facilitée par des hyperliens qui lui permettent de cliquer sur n'importe quel mot souligné afin de consulter le sens, notamment dans une phase de découverte et d'appropriation du lexique. La même activité peut être paramétrée sans cette facilitation lexicale dans une activité d'évaluation des acquis lexicaux en désactivant le filtre. Un glossaire peut, par ailleurs, être enrichi avec des éléments multimodaux (images, sons, vidéos) mais ChatGPT ne peut pas être convoqué dans une telle perspective car cela requiert un hébergement pérenne de ces éléments sur un serveur. Dans cette même perspective, nous avons voulu vérifier s'il était possible de générer un glossaire volumineux à partir de ressources lexicales brutes : pour ce faire, nous avons demandé à ChatGPT de nous produire un glossaire global intégrant tous les mots recensés par FleLex pour le niveau A1 et A2 et le résultat s'est avéré satisfaisant, couvrant l'ensemble des entrées disponibles.

3.3. Synthèse sur la génération des ressources et des activités

57 La mise en œuvre du protocole expérimental nous a permis de générer plusieurs ressources pédagogiques adaptées aux niveaux A1 et A2 du CECRL, compatibles pour une intégration dans Moodle. L'analyse des textes générés par ChatGPT montre que, malgré quelques ajustements lexicaux nécessaires, ces productions peuvent servir efficacement de documents déclencheurs. Les analyses lexicales effectuées à l'aide de FleLex révèlent que les textes incluent un mélange de mots adaptés aux différents niveaux, avec une majorité de mots fréquents généralement au niveau cible ou au niveau inférieur, ce qui justifie leur pertinence pour travailler sur le vocabulaire. L'utilisation de l'inventaire du CECRL (North & Fondation Eurocentres/Eaquals, 2015) permet de mieux cibler les tâches sociolangagières et le lexique des niveaux visés. Cet outil conjugué avec un prétraitement à l'aide de TreeTagger favorise la construction des activités de conceptualisation et de systématisation des faits de langue dans une visée d'appropriation de la grammaire. Les activités générées, bien que conformes aux formats Moodle, nécessitent, dans certains cas, un traitement ultérieur pour améliorer leur pertinence pédagogique, leurs consignes et leur compatibilité avec Moodle. La génération de glossaires s'est révélée être une tâche prometteuse car ChatGPT génère facilement des fichiers volumineux à partir d'un prétraitement à l'aide de TreeTagger. Les limites techniques des outils comme TreeTagger et les contraintes de ChatGPT pour les fichiers XML ont mis en lumière des pistes d'amélioration pour optimiser le processus. Globalement, les résultats confirment la fiabilité du protocole, tout en identifiant des ajustements nécessaires pour renforcer son efficacité dans une telle démarche et une confection précise et concise des lignes de commande afin d'éviter des résultats indésirables ou sujets à un traitement ultérieur qui risquerait de confondre l'utilisateur.

4. Conclusions et perspectives

58 Les résultats obtenus dans cette recherche montrent que la combinaison des outils de l'IAG, comme ChatGPT, et des technologies du

TAL comme FleLex et TreeTagger, permet de générer des ressources pédagogiques adaptées pour l'enseignement du FLE au niveau A1 et A2 du CECRL de manière rapide. Grâce à une supervision utilisant des documents spécifiques comme l'inventaire des contenus du CECRL (North & Fondation Eurocentres/Eaquals, 2015), il est possible d'assurer l'adéquation des ressources produites avec les niveaux visés. De manière similaire, la génération d'activités pédagogiques de grammaire et de compréhension écrite est également réalisable, bien que nécessitant une intervention humaine et pour garantir leur pertinence et leur adéquation avec les objectifs d'apprentissage. En revanche, l'intégration de ces ressources dans des plateformes comme Moodle reste limitée par la capacité de ChatGPT à produire de larges banques d'activités structurées en XML, ce qui impose des ajustements manuels ou des procédures fractionnées.

59 Cette exploration souligne ainsi l'utilité des outils du TAL, comme FleLex et TreeTagger, pour automatiser une partie du processus de conception pédagogique à l'aide de ChatGPT, notamment en matière d'analyse lexicale et morphosyntaxique. Cependant, la supervision humaine demeure indispensable pour corriger les imprécisions et garantir la qualité des productions. Au-delà de l'adéquation formelle ou linguistique des ressources, leur valeur pédagogique dépend aussi de leur intégration dans une unité didactique : il conviendrait, dans un prolongement de ce travail, d'envisager leur insertion dans des scénarios complets, articulant objectifs pédagogiques, tâches sociolangagières et modalités d'évaluation. Cela permettrait de mieux exploiter le potentiel des ressources générées et d'évaluer leur efficacité didactique à travers des indicateurs d'acquisition ou d'engagement.

60 Le protocole expérimental mis en place, bien qu'efficace, nécessite des ressources d'entraînement adaptées, ainsi qu'une réflexion méthodique sur la conception et la calibration des contenus, ce qui le rend chronophage pour certains utilisateurs non experts. Pour les experts, un protocole élargi peut être envisagé à travers l'introduction d'autres outils comme, par exemple, l'outil ReSyf – ressource lexicale pour l'étude des synonymes graduels en français (Billami & coll., 2018 ; Gala & Javourey-Drevet, 2020) ou, encore, des outils en ligne avec lesquels le corps enseignant en FLE ou dans d'autres langues est familier, ce qui demandera un recensement large des

pratiques pédagogiques en vigueur. Par ailleurs, il convient de rappeler que cette expérimentation a été réalisée avec le module gratuit de ChatGPT. En effet, il est nécessaire de réaliser le même protocole en utilisant d'autres modèles[18], ce que nous envisageons comme une des perspectives du présent travail. Dans cette optique, il nous semble pertinent de rappeler que les modèles de langue sont constamment nourris à partir de données régulièrement injectées ce qui pourrait, avec le temps, optimiser la capacité de ChatGPT à générer des contenus pédagogiques de plus en plus complexes et fiables. Cette étude confirme ainsi la pertinence d'une démarche de recherche-développement, qui permet d'articuler une génération automatique soumise à une double supervision – technique et pédagogique – à l'aide des outils du TAL et de l'inventaire du CECRL, qui peut être perfectible.

61 Enfin, ces résultats doivent être envisagés dans un contexte de constante évolution technologique, où les outils et les protocoles élaborés devront s'adapter à des environnements éducatifs en mutation constante. C'est pour cette raison que nous avons entrepris un premier test du protocole auprès d'une cohorte de 180 étudiants appartenant à deux promotions du master didactique des langues à l'Université de Montpellier Paul-Valéry dans le cadre de deux enseignements traitant des outils numériques pour l'enseignement des langues (M1) et de l'ingénierie pédagogique et de la formation (M2). Un deuxième test du protocole est envisagé auprès des concepteurs du réseau des Alliances françaises de la Bolivie et du Mexique dans le cadre du développement des modules DELF B2 et DALF C1 du projet Bolivia Habla Fr@ncés. Ces expérimentations visent également à permettre aux actuels et futurs enseignants de FLE de se saisir autant que faire se peut des différents outils issus de l'IAG et du TAL et d'œuvre au renforcement de leur littéracie numérique.

BIBLIOGRAPHIE

Albadarin, Y., Saqr, M., Pope, N., & Tukiainen, M. (2024). A systematic literature review of empirical research on ChatGPT in education. *Discover Education*, 3(1), 60. https://doi.org/10.1007/s44217-024-00138-2.

Allouche, A. (2023, octobre). *IA générative et ingénierie pédagogique : Le*

prompting, pistes de travail et applications. Carnet de veille « Éducation, numérique et recherche ». https://doi.org/10.58079/o3aq.

Antoniadis, G., Kraif, O., Lebarbé, T., Ponton, C., & Echinard, S. (2005). Modélisation de l'intégration de ressources TAL pour l'apprentissage des langues : La plateforme MIRTO. *Alsic*, 8(2), 65-79. https://doi.org/10.4000/alsic.376.

Billami, M. B., François, T. & Gala, N. (2018). ReSyf: A French lexicon with ranked synonyms. *Proceedings of the 27th International Conference on Computational Linguistics* (https://aclanthology.org/volumes/C18-1/), 2570-2581.

Bird, S., & Liberman, M. (2001). A formal framework for linguistic annotation. *Speech Communication*, 33(1-2), 23-60. https://doi.org/10.1016/S0167-6393(00)00068-6.

Bolivia habla fr@ncés—Plataforma pedagógica de las Alianzas francesas de Bolivia. (s. d.). Consulté 8 janvier 2025. https://boliviahablafrances.alianzafrancesa.org.bo/.

Bouillon, P. (1998). L'analyse et la génération. Dans P. Bouillon (dir.), *Traitement automatique des langues naturelles* (p. 27-33). De Boeck Supérieur. https://shs.cairn.info/traitement-automatique-des-langues-naturelles--9782801111819-page-27.

Burrows, A., Cetro, R., Miras, G. & Kübler, N. (2019). Présentation. *Éla. Études de linguistique appliquée*, 193(1), 5-8. https://doi.org/10.3917/ela.193.0005.

Burrows, A. & Miras, G. (2019). Pratiques numériques en langues : De la verbalisation à l'analyse des préoccupations enseignantes. *Alsic*, 22(2). https://doi.org/10.4000/alsic.3549.

Cavalla, C. (2020). La formation de futurs enseignants de FLE à la phraséologie et aux corpus numériques. *Phrasis-rivistadistudifraseologicieparemiologici*, 5, 27-40.

Conseil de l'Europe (2001). *Cadre européen commun de référence pour les langues : Apprendre, enseigner, évaluer*. Didier.

Danlos, L. (1991). Génération automatique de textes en langue naturelle. *Linx*, 4(1), 197-214. https://doi.org/10.3406/linx.1991.1198.

Danlos, L. & Roussarie, L. (2000). La génération automatique de textes. Dans J.-M. Pierrel (dir.), *Ingénierie des langues* (p. 311-330). Hermès.

Francois, T., Gala, N., Watrin, P. & Fairon, C. (2014). FLELex: A graded lexical resource for French foreign learners. *Proceedings of the Ninth International Conference on Language Resources and Evaluation* (LREC'14) (https://aclanthology.org/volumes/L14-1/), 3766–3773 http://www.lrec-conf.org/proceedings/lrec2014/pdf/1108_Paper.pdf.

Gala, N. & Javourey-Drevet, L. (2020). Mots « faciles » et mots « difficiles » dans ReSyf : un outil pour la didactique du lexique mobilisant polysémie, synonymie et complexité. *Lidil*, 62. https://doi.org/10.4000/lidil.8373.

Garcia-Debanc, C. (2001). L'étude des discours procéduraux aujourd'hui : Travaux linguistiques et psycholinguistiques. *Langages*, 35(141), 3-9. https://doi.org/10.3406/lgge.2001.871.

Garside, R., Leech, G. N., & McEnery, T. (dir.). (1997). *Corpus annotation : Linguistic information from computer text corpora*. Longman.

Glossaire – MoodleDocs. (s. d.). https://docs.moodle.org/4x/fr/Glossaire. Consulté le 6 mai 2024.

Granger, S. (2002). A Bird's-eye view of learner corpus research. Dans S. Granger, J. Hung & S. Petch-Tyson (dir.), *Language Learning & Language Teaching*, 6, 3-33). https://doi.org/10.1075/lllt.6.04gra.

Granger, S. (2007). Corpus d'apprenants, annotation d'erreurs et ALAO : Une synergie prometteuse. *Cahiers de lexicologie*, 2(91), 117-132. https://doi.org/10.48611/ISBN.978-2-8124-4343-5.P.0125.

Guichon, N. (2007). Recherche-développement et didactique des langues. *Recherches en didactique des langues et des cultures*, 4(4). https://doi.org/10.4000/rdlc.4903.

Guichon, N. (2012). *Vers l'intégration des TIC dans l'enseignement des langues*. Didier.

Guichon, N. (2015). Quelle transition numérique pour les étudiants internationaux ? *Alsic*, 18(1). https://doi.org/10.4000/alsic.2793.

Gustafson, K. L., & Branch, R. M. (avec ERIC Clearinghouse on Information & Technology & United States). (2002). *Survey of instructional development models* (4ème éd.). ERIC Clearinghouse on Information & Technology.

Higuera, C. de la & Iyer, J. (2024). *AI for Teachers: An Open Textbook*. https://pressbooks.pub/aiforteachers/.

North, B. & Fondation Eurocentres/Eaquals (2015). *Inventaire linguistique des contenus clés des niveaux du CECRL*. Eaquals. https://www.eaquals.org/wp-content/uploads/Inventaire_ONLINE_full.pdf.

Kaplan-Rakowski, R., Grotewold, K., Hartwick, P. & Papin, K. (2023). Generative AI and Teachers' Perspectives on Its Implementation in Education. *Journal of Interactive Learning Research*, 34(2), 313-338. https://doi.org/10.70725/815246mfssgp.

Kostov, J. (2020, mars 20). Moodle : Comment mettre en place une évaluation avec l'outil Test. Carnet de veille « NumeLang ». https://doi.org/10.58079/sgpe.

Liaison automatique des glossaires – MoodleDocs. https://docs.moodle.org/3x/fr/Liaison_automatique_des_glossaires. Consulté 9 janvier 2025.

Mayer, R. E. (2017). Using multimedia for e-learning. *Journal of Computer Assisted Learning*, 33(5), 403-423. https://doi.org/10.1111/jcal.12197.

Mayer, R. E. (2020). *Multimedia Learning* (3e éd.). Cambridge University Press; Cambridge Core. https://doi.org/10.1017/9781316941355.

Michard, A. (2000). XML : *Langage et applications*. Eyrolles.

Musial, M., & Tricot, A. (2020). *Précis d'ingénierie pédagogique*. De Boeck supérieur.

Nikolopoulou, K. (2024). Generative Artificial Intelligence in Higher Education: Exploring Ways of Harnessing Pedagogical Practices with the Assistance of ChatGPT. *International Journal of Changes in Education*, *1*(2), 103-111. https://doi.org/10.47852/bonviewIJCE42022489.

Ollivier, C. (2018). *Littératie numérique et approche socio-interactionnelle pour l'enseignement-apprentissage des langues*. https://www.ecml.at/Portals/1/5MTP/Ollivier/e-lang%20FR.pdf.

Ollivier, C., Gaillat, T., & Puren, L. (dir.). (2016). *Numérique et formation des enseignants de langue : Pistes et imaginaires*. Éditions des Archives contemporaines.

Pery-Woodley, M. P. (2001). Modes d'organisation et de signalisation dans des textes procéduraux. *Langages*, 35(141), 28-46. https://doi.org/10.3406/lgge.2001.873.

Pintard, A., & François, T. (2020). Combining Expert Knowledge with Frequency Information to Infer CEFR Levels for Words. Dans N. Gala & R. Wilkens (dir.), *Proceedings of the 1st Workshop on Tools and Resources to Empower People with REAding DIfficulties (READI)* (p. 85-92). European Language Resources Association. https://aclanthology.org/2020.readi-1.13/.

Poibeau, T. (2019). Le traitement automatique des langues : Tendances et enjeux. *Lalies*, 39, 7-65. https://doi.org/10.4000/1231f

Pôle d'élaboration des ressources linguistiques (PERL). perl-uspc. https://perl2018.wixsite.com/perl-uspc. Consulté 8 janvier 2025.

Ruggia, S. (2021). DeepFLE : L'intelligence artificielle pour prédire et décrire le(s) niveau(x) du CECRL d'un texte. *Didactique du FLES*, *2*(1), 103-109. https://doi.org/10.57086/dfles.253.

Ruggia, S., & Gaillat, T. (2023). Les corpus numériques pour la didactique des langues : De la formation des enseignants à l'élaboration de dispositifs d'apprentissage. *Corpus*, 24. https://doi.org/10.4000/corpus.8211.

Sauvage, J. (2019). Humanités numériques et formation innovante en didactique du FLE : Un enjeu pour une mobilité différenciée. *Recherches en didactique des langues et des cultures*, 16-2. https://doi.org/10.4000/rdlc.7051.

Schmid, H. (1997). Probabilistic part-of-speech tagging using decision trees. Dans D. B. Jones & H. Somers (dir.), *New Methods In Language Processing* (p. 154-164). Routledge. https://doi.org/10.4324/9781315072685.

Schmid, H. (1999). Improvements in Part-of-Speech Tagging with an Application to German (Vol.11). Dans S. Armstrong, K. Church, P. Isabelle, S. Manzi, E. Tzoukermann, & D. Yarowsky (dir.), *Natural Language Processing Using Very Large Corpora* (p. 13-25). Springer Netherlands. https://doi.org/10.1007/978-94-017-2390-9_2.

Schmid, H. & Laws, F. (2008). Estimation of Conditional Probabilities With Decision Trees and an Application to Fine-Grained POS Tagging. *Proceedings*

of the 22nd International Conference on Computational Linguistics (Coling 2008) (p. 777-784). https://aclanthology.org/C08-1098.

Sébillot, P. (2015). Chapitre 3. Le traitement automatique des langues face aux données textuelles volumineuses et potentiellement dégradées : Qu'est-ce que cela change ? Dans *Big Data* (p. 43-60). De Boeck Supérieur. https://shs.cairn.info/big-data--9782804189150-page-43.

Son, J.-B., Ružić, N. K, & Philpott, A. (2023). Artificial intelligence technologies and applications for language learning and teaching. *Journal of China Computer-Assisted Language Learning*, 0(0). https://doi.org/10.1515/jccall-2023-0015.

Tagset FR. https://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger/data/french-tagset.html. Consulté 8 mai 2024.

Test – MoodleDocs. https://docs.moodle.org/4x/fr/Test. Consulté 6 mai 2024.

TreeTagger. https://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger/. Consulté 7 mai 2024.

TreeTagger – Centre de traitement automatique du langage. https://cental.uclouvain.be/treetagger/. Consulté 7 mai 2024.

Tovar, L., Stefanou, D. & Kostov, J. (2023, novembre). *Hybrider pour mieux accompagner en LANSAD : l'exemple du projet Nexus de l'université Paul-Valéry Montpellier 3*. Communication présentée à RANACLES 2023 – Congrès du rassemblement national des centres de langues de l'Enseignement supérieur, École nationale supérieure de Lyon, Lyon, France.

UNSECO (2019). *Beijing Consensus on Artificial Intelligence and Education*. UNESCO Nations Educational, Scientific and Cultural Organization.

Zock, M., & Sabah, G. (1992). La génération automatique de textes : trente ans déjà, ou presque. *Langages*, 26(106), 8-35. https://doi.org/10.3406/lgge.1992.1630.

NOTES

1 Le présent article se fonde sur une expérimentation à partir de la version gratuite de l'outil, accessible à n'importe quel internaute.

2 https://cental.uclouvain.be/cefrlex/flelex/ (dernière consultation le 20 décembre 2024).

3 https://cental.uclouvain.be/treetagger/ (dernière consultation le 20 décembre 2024).

4 Connue également sous la dénomination « approche symbolique ».

5 https://www.ai4t.eu/ (dernière consultation le 20 décembre 2024).

6 https://eduscol.education.fr/692/les-travaux-academiques-mutualises-traam (dernière consultation le 20 décembre 2024).

7 https://pedagogie.ac-strasbourg.fr/lettres/traam-2024/ (dernière consultation le 20 décembre 2024).

8 https://h5p.org/ (dernière consultation le 20 décembre 2024).

9 https://moodle.unifr.ch/course/view.php?id=280393 (dernière consultation le 25 janvier 2025).

10 https://perl2018.wixsite.com/ (dernière consultation le 20 décembre 2024).

11 https://boliviahablafrances.alianzafrancesa.org.bo/ (dernière consultation le 20 décembre 2024).

12 Abréviation de « Analyse, design, développement, implémentation, évaluation ».

13 Voir https://www.liquid-technologies.com/online-xml-validator (dernière consultation le 15 février 2025).

14 Appelée également un *prompt*, la ligne de commande est un moyen d'interagir avec un système informatique à l'aide d'instructions très courtes. Dans le domaine de l'IAG, il existe un domaine de recherche spécifique aux *prompts* (le « *prompting* ») : leur clarté et concision sont gageures d'un travail de génération efficace et efficient. Pour plus d'informations sur le fonctionnement et les caractéristiques des prompts, nous renvoyons le lecteur vers le blog d'Allouche qui contient une pléthore de travaux de recherche et des références sur ce sujet : https://edunumrech.hypotheses.org/ (dernière consultation, le 20 décembre 2024).

15 Questions à réponse unique / questions à réponses multiples.

16 Une typologie relativement complète des exercices Moodle peut être consultée à l'adresse https://doi.org/10.58079/sgpe (dernière consultation, le 20 décembre 2024). Cette typologie varie d'une plateforme à une autre : Moodle est livré avec un nombre restreint d'outils et de plugins que les administrateurs peuvent enrichir en fonction des besoins des utilisateurs et, très souvent, suivant les politiques des établissements et des structures qui gèrent les LMS. C'est pour cette raison que l'utilisateur doit avoir au préalable une vision d'ensemble des types d'exercices existants sur la plateforme dans laquelle il effectuera l'importation.

17 https://docs.moodle.org/4x/fr/Liaison_automatique_des_glossaires (dernière consultation le 25 janvier 2025).

18 L'outil fonctionne en utilisant des modèles de langue différents que l'on peut consulter dans la documentation officielle : https://platform.openai.com/docs/models (dernière consultation, le 25 février 2025).

RÉSUMÉS

Français

Cet article examine la synergie entre les outils d'intelligence artificielle générative (IAG), les technologies du traitement automatique des langues (TAL) et les plateformes numériques d'enseignement (LMS) pour la conception de ressources pédagogiques destinées à l'enseignement du français langue étrangère (FLE), particulièrement aux niveaux A1 et A2 du CECRL. Il s'interroge sur la capacité de ces outils, et notamment ChatGPT, FleLex et TreeTagger, à produire des contenus pédagogiques conformes aux descripteurs du CECRL tout en répondant aux exigences des environnements numériques visés.

Après une introduction exposant les fondements théoriques et les questions de recherche, la deuxième section propose un état de l'art détaillé. Elle retrace l'évolution des technologies de génération automatique en TAL et des recherches sur l'intelligence artificielle appliquée à l'éducation, en les mettant en relation avec les défis actuels de la didactique des langues et les besoins des enseignants. La troisième section décrit les outils mobilisés et le protocole expérimental. Une méthodologie combinant une approche non supervisée et supervisée à l'aide des outils TAL est mise en œuvre pour produire et calibrer des ressources pédagogiques. Les ressources et les activités générées incluent des textes, des images, des activités de compréhension écrite, de grammaire et des glossaires interactifs, le tout intégré dans Moodle via des structures XML adaptées. La quatrième section est consacrée à l'analyse des ressources générées. Elle explore leur qualité formelle et leur conformité avec les niveaux du CECRL et leur intégration dans des activités pédagogiques structurées. Des perspectives d'amélioration pour une utilisation optimisée des outils sont également identifiées.

L'article se conclut par une réflexion sur les implications pédagogiques et les perspectives futures, tout en envisageant des pistes pour la formation des enseignants et pour l'optimisation des outils pour mieux intégrer ces technologies dans des pratiques éducatives durables.

English

This paper examines the synergy between generative artificial intelligence (AI), natural language processing (NLP) technologies, and learning management systems (LMS) in the design of pedagogical resources for teaching

French as a Foreign Language (FFL), particularly at the Common European Framework of Reference for Languages (CEFR) levels A1 and A2. It investigates the potential of tools such as ChatGPT, FleLex and TreeTagger, to produce educational content aligned with CEFR descriptors while meeting the requirements of targeted digital learning environments.
After an introduction outlining the theoretical foundations and research questions, the second section provides a detailed state of the art. It traces the evolution of automatic text generation technologies in NLP and research on AI applied to education, linking them to current challenges in language teaching and teachers' needs. The third section describes the tools used and the experimental protocol. A methodology combining unsupervised and supervised approaches with NLP tools is implemented to create and calibrate pedagogical resources. The generated resources and activities include texts, images, reading comprehension exercises, grammar activities, and interactive glossaries, all integrated into Moodle through adapted XML structures. The fourth section focuses on analyzing the generated resources. It explores their formal quality, alignment with CEFR levels, and integration into structured pedagogical activities. Suggestions for optimizing the tools for improved usability are also identified.
The article concludes with a reflection on pedagogical implications and future perspectives, while considering pathways for teacher training and the optimization of tools to better integrate these technologies into sustainable educational practices.

INDEX

Mots-clés
français langue étrangère, génération automatique de textes, intelligence artificielle générative, traitement automatique des langues, Moodle

Keywords
french as a foreign language, natural language generation, generative artificial intelligence, natural language processing, Moodle

Rubriques
Recherche et pratique

AUTEUR

Jovan Kostov
Jovan Kostov est maître de conférences en sciences du langage et en didactique du FLE à l'université de Montpellier Paul-Valéry. Ses recherches portent sur l'ingénierie pédagogique et la formation en langue ainsi que leur modélisation à

travers le TAL, la linguistique de corpus et les humanités numériques. Il coordonne plusieurs projets de recherche-action internationaux sur les dispositifs numériques d'enseignement/apprentissage du FLE dans un contexte latino-américain et européen.
jovan.kostov@univ-montp3.fr
IDREF : https://www.idref.fr/195813154
ORCID : http://orcid.org/0000-0003-3586-0273
HAL : https://cv.archives-ouvertes.fr/jovan-kostov