

Les textes générés par l'IA comme modèle pédagogique en classe de FLE: aspects théoriques et pratiques

Sara COTELLI KURETH

Hasti NOGRECHI

Université de Neuchâtel

Institut de langue et civilisation françaises

Fbg. de l'Hôpital 61-63, 2000 Neuchâtel, Suisse

sara.cotelli@unine.ch

ORCID: 0000-0002-5213-2801

hasti.noghrechi@unine.ch

ORCID: 0000-0002-9494-4664

New generative AI (GenAI) tools, machine translation and automatic text production (ATP) tools like ChatGPT are being used more and more in language classrooms, by teachers and learners alike. Integrating the output of these systems as a model for students raises several questions, some of which we discuss and address in this article. We think the way that data-driven learning was successfully introduced in language learning and teaching (LLT) is an interesting basis to consider the introduction of GenAI tools in the classroom. The first part of this text discusses theoretical principles linked to the use of these technologies by language learners. The second part presents an action-research study in which we compared the use of an ATP tool (Copilot) to the use of a corpus (Corpus Leipzig - French) in a B1 French as a foreign language vocabulary class. Students were divided in two groups and each group used one tool to find collocations for a set of words. The results show no significant difference in the way participants evaluated the usefulness of the tool and we found that the two groups gathered a similar number of correct items. However, the level of difficulty of the words selected by participants differed between the two groups: Copilot provided more advanced vocabulary (B2 and above) and students choosing from the corpus tended to pick words below or at their level. In conclusion, it seems that for this task ATP tools can successfully be used but do not offer any real added value to other tools that have been widely available for some time, namely corpora.

Keywords:

French as a foreign language teaching; AI; automatic text production tools; collocations; data-driven learning.

Mots-clés:

FLE; IA; générateurs automatiques de textes (GAT); collocations; enseignement à partir des corpus.

1. Introduction¹

L'arrivée de l'intelligence artificielle générative (IAGen), notamment la traduction automatique neuronale (TA) et les générateurs automatiques de textes (GAT) comme ChatGPT, a ouvert un nouveau champ de réflexion pour l'enseignement/apprentissage des langues étrangères (LE). L'intégration positive de la TA dans l'enseignement a été démontrée par de nombreuses

¹ Nous voulons remercier les deux personnes anonymes qui ont relu notre article. Grâce à leur relecture critique et bienveillante, nous avons pu améliorer cette version du texte.

recherches (voir la synthèse proposée par Jolley & Maimone 2022 et Klimova et al. 2023 et Raaflaub et Reber 2025 dans ce volume), notamment pour favoriser l'autonomie des apprenant·es (Cotelli Kureth & Noghrechi 2023).

La démocratisation et la normalisation de l'usage de l'IA dans l'éducation ces dernières années soulèvent des enjeux moraux et éthiques (Adams et al. 2023), interrogent les pratiques des enseignant·es (Yanuha 2020) et l'impact de ces outils sur l'enseignement/apprentissage des LE (Motlagh et al. 2023), notamment dans la rédaction de textes académiques (Crompton & Burke 2023).

Les utilisations de l'IAGen dans l'apprentissage des LE sont multiples: l'output de l'IA peut servir de modèle, de support pour la rédaction d'un texte ou pour une tâche d'expression orale. Les apprenant·es peuvent utiliser ChatGPT comme un·e tuteur·trice, pour poser des questions ou pour obtenir un feedback, ou comme un partenaire pour pratiquer l'expression orale (Molle & Hamza-Jamann 2025). Selon nos discussions avec des collègues européen·es, on constate en outre une utilisation croissante de ces outils par les enseignant·es pour produire des textes 'sur mesure' ou pour transformer des textes trouvés sur Internet (les résumer, paraphraser ou simplifier) (voir aussi Galindo-Dominguez et al. 2023). De nombreuses applications spécialisées offrent ce service gratuitement d'un simple clic (voir par ex. Twee, <https://twee.com/> ou brisk teaching, <https://www.briskteaching.com/>). Plusieurs gouvernements et institutions ont créé des guides pour présenter les principes éthiques à respecter lors de l'intégration de l'IA dans l'éducation (voir par exemple Gouvernement du Québec 2024)

Cette banalisation du produit de l'IAGen comme modèle pour les apprenant·es soulève certaines interrogations auxquelles nous envisageons de répondre dans cet article au travers du prisme du *data-driven learning* (DDL). Nous chercherons ainsi à savoir dans quelle mesure les outils d'IAGen constituent une alternative efficace aux corpus dans l'enseignement du FLE. La première partie rappelle certains points importants du DDL, décrit le fonctionnement des grands modèles de langage (LLM), et analyse les limites des textes produits par l'IAGen. La seconde présente une recherche-action menée dans une classe de FLE de niveau B1, sur une tâche de recherche de cooccurrences avec un corpus et avec un GAT. Nous explorons les attitudes des apprenant·es face aux deux outils et les résultats de leurs recherches pour évaluer la pertinence d'utiliser un GAT pour la recherche de cooccurrences.

2. Inclusion de l'IAGen comme modèle pour les apprenant·es: quelques considérations théoriques

2.1 Data-driven learning ou apprentissage sur corpus

Pour mieux comprendre le "data-driven learning" (Johns & King 1991) ou "apprentissage sur corpus" (Tyne 2013; Boulton & Tyne 2014), il faut d'abord

définir la notion de corpus qui est comprise différemment selon les auteur·es (Boulton et al. 2015). Les linguistes se basent fréquemment sur une définition (voir notamment Gilquin & Gries 2009) qui présente le corpus comme "un recueil électronique d'un grand nombre de textes (écrits comme oraux) pouvant être interrogé à l'aide d'un logiciel dédié (ou concordancier)" (Leray & Tyne 2016: 132).

L'apprentissage sur corpus consiste à "using the tools and techniques of corpus linguistics for pedagogical purposes" (Gilquin & Granger 2010: 1). C'est une approche d'enseignement des langues basée sur les données sélectionnées dans les corpus. Dans cette méthode, les apprenant·es sont actif·ves et découvrent différents aspects langagiers par eux·elles-mêmes en examinant des extraits de données provenant de corpus (oraux ou écrits). Iels adoptent donc une approche déductive dans leur apprentissage (Gilquin & Granger 2010). Selon Johns (1997: 101): "every student [is] a Sherlock Homes". Ainsi, pour Mair (2002), cette approche augmente considérablement leur motivation. Le rôle de l'enseignant·e change également: iel n'est plus considéré comme la seule source d'information (Brown et al. 2007) et son travail en classe se limite à guider les apprenant·es dans la bonne direction (Leray & Tyne 2016).

De surcroît, l'emploi de cette approche dans l'enseignement des LE donne accès aux documents authentiques en classe (Boulton & Tyne 2014). Elle permet également aux apprenant·es de corriger leurs productions orales ou écrites (Nesselhauf 2004) mais elle développe aussi les compétences métalinguistiques des apprenant·es, comme par exemple développer leur compétences à reconnaître des particularités langagières (O'Sullivan 2007: 278).

Malgré tous ces avantages, l'utilisation de ces outils peut avoir certains inconvénients, ce qui explique la méfiance d'une partie du corps enseignant. Certaines de ces difficultés sont plutôt de type langagier: le lexique peut être inadapté ou les structures syntaxiques trop complexes et non attendues pour le niveau des apprenant·es (Richards 2009). On rencontre aussi des problèmes de compréhension chez les apprenant·es débutants (Thomas et al. 2016). Finalement, certaines représentations culturelles très fortes freinent parfois les effets positifs de cette méthode d'enseignement (Boulton 2009, 2017). En effet, le DDL redéfinit les rôles traditionnellement attribués aux acteurs·trices dans le domaine didactique, les enseignant·es peuvent craindre une perte de contrôle ou de statut (voir ci-dessus, ainsi que Brown et al. 2007; Leray & Tyne 2016), tandis que les apprenant·es peuvent être réticent·es à assumer une plus grande responsabilité dans leur apprentissage (Boulton 2007: 38).

Dans cette perspective, il paraît intéressant de se pencher sur l'utilisation de l'IAGen et plus particulièrement des GAT, qui s'appuient eux-aussi sur une recherche statistique dans de vastes corpus. Il nous semble possible d'intégrer l'IAGen d'une façon similaire au DDL. D'ailleurs, ces nouveaux outils ne

permettraient-ils pas de pallier ces points négatifs de l'approche de l'apprentissage sur corpus, tout en conservant les points forts de celle-ci?

2.2 L'IA Gen: une meilleure solution?

2.2.1 Quelques considérations techniques

Nous ne décrivons pas ici en détail le fonctionnement de l'IA Gen (voir Benites et al. 2023, Cain et al. 2023, Sarrion 2023 pour plus d'informations). Nous entendons par IA Gen, tout système d'IA qui génère du texte, de l'image ou du son. Cet article se concentre sur les générateurs de textes basés sur un prompt, comme ChatGPT, Gemini ou Copilot et nous nous intéressons particulièrement aux éléments techniques qui impactent le plus leur utilisation. Nous souhaitons tout de même mentionner brièvement en introduction trois points de tension éthiques essentiels à rappeler dans toute discussion sur l'IA Gen:

- 1. les problèmes suscités par la sécurité des données, que ce soit l'utilisation de matériel sous copyright pour l'entraînement des systèmes ou la conservation par les systèmes de tous les textes que nous leur soumettons pour des besoins d'entraînement;
- 2. les problèmes en lien avec l'intégrité scientifique qui font couler beaucoup d'encre et divisent la société scientifique;
- et 3. l'impact environnemental de l'entraînement et de l'utilisation de l'IA Gen, qui en fait un outil peu durable.

Avec ces points en tête, passons en revue les deux technologies essentielles sur lesquels reposent ces outils.

La première caractéristique technologique est celle des grands modèles de langage (LMM). Ils sont constitués d'énormes recueils de textes qui rassemblent des centaines de millions de documents plutôt écrits (livres, articles scientifiques, textes provenant d'Internet, etc.). Les entreprises développant les principaux outils d'IA ne sont pas transparentes quant au contenu exact de ces énormes recueils. La dimension et le contenu du LLM sur lequel le système est entraîné a pourtant une incidence directe sur la qualité de son produit. Ainsi, les langues à grande utilisation, comme l'anglais et certaines langues européennes, disposent de corpus de textes numérisés très importants et les outils d'IA sont donc performants pour ces langues. Pour d'autres qui sont moins numérisées, les systèmes actuels ne sont pas optimaux. Il s'agit souvent des langues premières des migrant·es que nous accueillons dans nos cours de FLE (par ex. l'albanais, l'arabe, le turc, l'ukrainien, etc.). Ces étudiant·es sont donc désavantagé·es par rapport aux locuteurs·trices de langues à grande utilisation, ce qui crée des inégalités.

Deuxièmement, les GAT génèrent leurs textes grâce à des calculs statistiques, selon un modèle de prédiction du mot suivant. Il est possible pour les concepteurs·trices, et parfois pour les utilisateurs·trices, de régler les outils pour qu'ils produisent des textes plutôt stéréotypés ou plus créatifs. Cette

technologie permet de générer des textes qui sont d'une très bonne qualité linguistique et idiomatique.

Produisant un résultat basé sur une recherche statistique dans d'énormes recueils de textes authentiques², les outils d'IAGen semblent donc intéressants pour une utilisation en classe de LE. Cette technologie permet au produit de l'IAGen de présenter des mots, des tournures syntaxiques, des collocations qui sont courantes et apparaissent souvent dans les corpus, et qui sont ainsi utiles aux apprenant·es. Les essais avec la TA dont la technologie repose également sur les LLM et sur des algorithmes statistiques a montré son utilité pour les apprenant·es qui l'utilisent comme un texte modèle pour corriger leur propre brouillon ou préparer un script pour l'expression orale (Tsai 2019; Lee & Briggs 2021). Nous l'avons testé avec succès pour la découverte de règles grammaticales avec des étudiant·es A1 (Cotelli Kureth & Nogrechi 2024). Toutefois, les recherches menées sur la TA montrent également certains problèmes des textes produits par ces technologies. Qu'en est-il de l'output des GAT?

2.2.2 Quels problèmes avec le produit de l'IAGen?

Il existe peu de données basées sur la recherche qui concernent directement les GAT, vu leur relative nouveauté. Il est toutefois possible d'inférer certains points des nombreuses recherches qui ont été effectuées sur le produit de la TA (Benites et al. 2023). On peut résumer les problèmes rencontrés lorsqu'on produit un texte avec l'IAGen en deux points principaux: les faiblesses de la machine et les faiblesses des (re)lecteurs·trices.

Les biais que présente le produit de la TA sont un des points faibles les plus discutés et les plus analysés, notamment les biais de genre (Savoldi et al. 2021), syntaxiques (Hovy et al. 2020) et culturels (Ji 2023), mais aussi le capacitisme (Primerano 2022). Dans le but de conserver une atmosphère inclusive dans les classes de LE, les enseignant·es doivent donc porter une attention particulière aux textes rédigés par l'IAGen pour s'assurer de les contrer.

Ensuite, il a été démontré que les textes produits par la TA utilisent une "artificially impoverished language" (Vanmassenhove et al. 2021: 2203), parfois nommée "machine translationese" (Loock 2020). Peu de recherches ont jusqu'ici été effectuées sur le produit de la GAT, mais les premiers résultats tendent à montrer que celui-ci a beaucoup de similitudes avec le produit de la TA: petit stock de connecteurs souvent répétés (Shaitarova 2023), moins de complexité syntaxique et de pronoms que dans les textes rédigés par les êtres humains (Kurokawa & Salingre 2025). Ce manque de créativité des textes rédigés par l'IA pourrait poser des problèmes aux apprenant·es plus avancés.

² Nous comprenons "authentique" dans le sens d'"authentique" en enseignement/ apprentissage des LE, c'est-à-dire un document qui n'a pas été créé à des fins didactiques (Boulton 2009).

Toutefois, le vocabulaire et la syntaxe utilisés par les GAT vont bien au-delà des connaissances des débutant·es et l'IA permet de leur proposer des modèles intéressants. En outre, la répétition de certaines structures peut même leur être bénéfique et permettre une meilleure mémorisation, voir développer l'apprentissage fortuit (incidental learning, voir Hulstijn 2013).

Un dernier point problématique des textes produits par l'IAGen est les erreurs occasionnelles qu'ils contiennent encore, malgré leur excellente qualité. Celles-ci sont souvent difficiles à déceler, mais peuvent produire des ambiguïtés et des énoncés illogiques (Shi et al. 2022). Pour la TA, les erreurs portent surtout sur des questions sémantiques: le texte-cible ne dit pas tout à fait la même chose que le texte-source. On trouve un problème similaire avec les GAT lorsqu'on leur demande de reformuler, résumer ou corriger un texte: le produit d'arrivée propose parfois un contenu quelque peu différent. C'est pour cette raison, qu'il est essentiel de toujours réviser avec beaucoup d'attention les textes produits par l'IAGen, non seulement du point de vue de la correction linguistique, mais en se focalisant sur le sens. Cette révision pose d'ailleurs parfois des problèmes en raison de certaines fragilités des (re)lecteurs·trices humain·es, parmi lesquelles on peut compter le minimalisme dont iels font souvent preuve en interrogeant l'IAGen.

Finalement, grâce à des études menées auprès de traducteurs·trices professionnel·les (Wang & Daghigh 2023), nous savons que travailler avec des textes produits par l'IA nécessite un niveau de concentration élevé, en raison de ce que l'on appelle la *fausse aisance* de ces textes ("false fluency", voir Martindale & Carpuat 2028). Ce concept fait référence à la surface linguistique parfaite des textes produits par les systèmes de TA, derrière laquelle se cachent parfois un contenu sémantiquement incorrect qui est souvent difficile à détecter (Loock & Léchaugnette 2021). Ce travail est bien sûr encore plus délicat pour les personnes qui rédigent dans une L2 (Holt et al. 2022). Par conséquent, si on utilise la TA ou un GAT pour rédiger un texte dans une L2, il est fort probable que la fluidité du texte produit par ces systèmes nous incite à adopter un contenu qui n'exprime pas tout à fait ce que nous avons initialement prévu.

2.3 Synthèse: le produit de l'IAGen pour la classe de FLE

Au final, il nous apparaît pertinent d'utiliser le produit de l'IAGen comme modèle pour les apprenant·es de LE, mais un réservant quelques aménagements.

En raison de la nature même des LLM et de la technologie basée sur les probabilités statistiques, le contenu grammatical et lexical des textes générés par ces systèmes est optimal pour les apprenant·es débutant·es à moyen·nes. La simplicité de la syntaxe et le vocabulaire moins varié ont en effet peu d'impact pour ce public et, même, la répétition de certaines structures, connecteurs ou autre peut être un atout. Les étudiant·es plus avancé·es pourront profiter de travailler à partir de ces textes, mais avec une vision plus

critique, en tentant de les améliorer, de les transformer pour d'autres contextes communicatifs, par exemple. La première question que l'enseignant·e doit se poser est donc celle du public et de la tâche.

Ensuite, les enseignant·es qui introduisent des textes rédigés grâce à l'IAGen doivent toujours en faire une révision approfondie pour corriger d'éventuelles fautes, vérifier que le sens initial a été conservé et contrer des biais. Cela introduit une contrainte à l'utilisation de cette technologie à laquelle on fait souvent appel pour gagner du temps. Il faut toutefois reconnaître que lorsqu'il est utilisé de façon critique, cet outil n'est pas forcément un raccourci.

Il est finalement important d'accompagner les apprenant·es dans leur utilisation des GAT. Eux·elles-mêmes ont besoin de connaissances de base sur le fonctionnement de l'outil, ses faiblesses et leurs faiblesses pour en faire un allié dans leur processus d'apprentissage. Introduire le produit de l'IAGen en classe permet ainsi à l'enseignant·e de développer pour les apprenant·es une littéracie propre à ces outils (Cardon et al. 2023; Cotelli Kureth et al. 2025).

Les résultats théoriques soulèvent des questions intéressantes quant à l'efficacité pratique des outils d'IAGen en classe de FLE. Nous allons examiner empiriquement un exemple d'introduction d'un GAT dans une tâche d'apprentissage dans la section suivante. Nous allons analyser la pertinence de cet outil pour la recherche de cooccurrences, sur le modèle d'expériences similaires avec le DDL.

3. Corpus et IAGen pour la recherche de cooccurrences

La recherche-action que nous présentons dans cet article a été intégrée dans un cours de vocabulaire de niveau B1, dans le cadre de l'année propédeutique en FLE à l'Institut de langue et de civilisation françaises de l'Université de Neuchâtel. Les auteures ont développé une tâche de recherche de cooccurrences avec les 73 étudiant·es de cette filière. La classe a été divisée en deux (un groupe de 37 et un groupe de 36 étudiant·es) et chaque moitié a réalisé le même exercice avec un outil différent: un groupe a fait ses recherches avec le Corpus français Leipzig (CFL)³ et l'autre groupe avec le GAT disponible par leur compte informatique à l'université, Copilot de Microsoft. Nous avons choisi le CFL (voir Kamber 2011) pour effectuer cette activité de DDL avec nos étudiant·es car il est couramment utilisé par plusieurs enseignant·es de l'ILCF et il est relativement facile d'accès.

Dans le cadre de l'apprentissage du vocabulaire, l'importance de l'enseignement des cooccurrences a souvent été soulevée (voir Cavalla 2019). Une tâche pour conscientiser ce phénomène avait donc toute sa place dans ce cours. De plus, plusieurs chercheurs·euses ont montré l'intérêt du DDL pour

³ URL: https://corpora.uni-leipzig.de/fr?corpusId=fra_mixed_2012

l'enseignement/apprentissage de ces 'chunks' langagiers (par ex. Kamber 2011; Sun & Park 2023), ce qui a encore renforcé la pertinence de cette tâche pour notre recherche-action. On peut définir la *cooccurrence* comme "une combinaison de deux mots qui apparaissent fréquemment ensemble dans des énoncés" (Anctil & Tremblay 2016: sp). Dans certains cas, un des deux mots perd son sens premier et les linguistes nomment ces cooccurrences des *collocations*. Nous n'avons pas fait cette distinction terminologique auprès de nos étudiant·es et avons parlé pour les deux cas de *cooccurrences*.

Cette recherche-action voulait répondre à la question de recherche suivante: est-ce que les deux outils permettent un résultat comparable? C'est-à-dire, est-ce que les apprenant·es obtiennent les mêmes collocations lorsqu'ils cherchent avec un corpus ou avec un GAT? Un des outils est-il plus efficace ou plus facile à utiliser pour la recherche de cooccurrences pour des apprenant·es B1?

3.1 Contexte de la recherche-action

Pour la tâche au cœur de notre recherche-action, les participant·es ont été invité·es, selon leur groupe, à rechercher, à l'aide soit d'un corpus, soit d'un GAT, une liste de verbes coocurents à 8 noms sélectionnés.

Le choix des noms a été fait à partir des mots les plus fréquents trouvés dans un corpus constitué sur AntConc⁴, une suite logicielle gratuite pour l'analyse de corpus, la concordance et l'analyse de texte. Nous avons collecté tous les textes intégrés aux différents cours du semestre d'automne de l'année propédeutique (utilisés dans les cours de compréhension écrite, d'expression écrite, de vocabulaire, d'oral en situation, de lecture à haute voix, d'orthographe, de grammaire, etc.)⁵, qui ont été fournis aux auteures par leurs collègues. Le corpus total comprenait 44 documents et 37'584 tokens. Après une recherche KWIC, nous avons choisi les noms le plus mentionnés dans le corpus qui présentaient un nombre élevé de cooccurrences avec des verbes dans les dictionnaires consultés⁶: *métier* (52 mentions dans le corpus), *enfant* (50), *travail* (41), *besoin* (39), *place* (37), *temps* (37), *émotion* (34) et *œil* (33).

Lors de la leçon, chaque groupe a été accompagné par une des auteures. La leçon a commencé par une brève définition et quelques explications sur la notion de cooccurrence. Chaque groupe a ensuite découvert l'outil qu'il allait utiliser. Dans le cas du groupe Copilot, une brève initiation à la rédaction de

⁴ URL: <https://www.laurenceanthony.net/software/antconc/>

⁵ Le cursus de l'année propédeutique a été construit en commun entre les enseignantes qui partagent les mêmes textes pour les différents cours, dans le but d'améliorer l'apprentissage des étudiant·es.

⁶ En particulier Le dictionnaire des cooccurrences en ligne, proposé sur le site du Bureau de la traduction du Gouvernement du Canada (URL: <https://www.noslangues-ourlanguages.gc.ca/fr/dictionnaire-des-cooccurrences/index-fra#mots-cles>) et le *Dictionnaire des combinaisons de mots*, publié par Le Robert (2007).

prompts a mis l'accent sur trois aspects: l'importance de donner du contexte, proposer un rôle à la machine, et donner suffisamment de détails sur le produit attendu. Ensuite, les participant·es ont dû compléter une phrase où le verbe manquait, grâce aux informations trouvées, suivant le groupe, soit dans le corpus, soit grâce à Copilot,: "Je un métier que j'adore. J'aime parler et rencontrer des gens, donc c'est parfait pour moi d'être réceptionniste". Nous avons utilisé l'exemple de prompt suivant:

Je suis un étudiant / une étudiante de français de niveau B1 en Suisse. Ma langue maternelle est le xxx et j'aimerais améliorer mon expression écrite. Mon enseignante m'a demandé de trouver des collocations avec le mot métier, comme par exemple dans la phrase "Je un métier que j'adore. J'aime parler et rencontrer des gens, donc c'est parfait pour moi d'être réceptionniste". Donne-moi les verbes qui peuvent être utilisés dans la phrase. Ensuite, donne-moi 10 autres verbes qu'on peut utiliser avec métier comme sujet et 10 verbes avec métier comme objet. Donne-moi une phrase d'exemple pour chaque verbe.

Cette première étape s'est déroulée en plenum. Ensuite, iels devaient chercher d'autres verbes et former des phrases d'exemple, en autonomie. Après ce travail en autonomie, iels échangeaient avec leurs voisin·es. Lors de la leçon, les étudiant·es ont travaillé sur deux noms (*métier* et *émotion*), puis iels ont travaillé sur les six noms restants en devoirs. Durant le cours, les participant·es pouvaient solliciter de l'aide auprès de l'enseignante qui a pu les guider. Lors du cours suivant, l'enseignante a repris dans un exercice une partie des verbes recueillis par les étudiant·es pour proposer une courte liste de verbes intéressants à apprendre et pour que tous·tes les apprenant·es aient pu profiter équitablement de la séquence didactique.

3.2 Données et méthodologie

Pour répondre à nos questions de recherche, nous avons recueilli deux types de données:

1. Les verbes et les phrases d'exemple récoltées par les apprenant·es lors de la leçon et pour les devoirs. Il faut noter qu'iels n'ont pas toutes et tous soumis les devoirs, mais une petite majorité les a faits (45/73).
2. Les participant·es ont répondu lors du cours suivant à un questionnaire créé sur Qualtrics⁷ qui leur posait huit questions à choix multiples et une question ouverte pour tester leur avis sur la tâche et l'outil utilisé. Nous voulions savoir si les participant·es allaient dans le futur utiliser l'outil découvert pour rechercher d'autres cooccurrences et si iels allaient l'utiliser dans leur apprentissage du français. Les autres questions cherchaient à connaître leur avis sur l'outil avec lequel iels avaient effectué leurs recherches. Était-il: facile à utiliser, pratique pour trouver des résultats, bon pour l'apprentissage des cooccurrences?

⁷ <https://www.qualtrics.com/>

Les verbes récoltés par les apprenant·es ont été comptés et comparés à la liste des verbes proposés dans le *Dictionnaire des combinaisons de mots* (Le Robert, 2007) et le dictionnaire des cooccurrences en ligne, proposée sur le site du Bureau de la traduction du Gouvernement du Canada. Ces deux dictionnaires ont été choisis pour leur facilité d'accès pour nos étudiant·es, le premier étant disponible à la bibliothèque de l'ILCF et le second en libre accès en ligne. Nous avons donc considéré, de façon arbitraire, que seuls les verbes qui se trouvaient sur cette liste étaient corrects. Nous avons ensuite comparé statistiquement les résultats obtenus par le groupe Corpus Français Leipzig (CFL) et le groupe Copilot en termes de nombre de verbes qui étaient proposés et de nombre de verbes corrects. Un modèle de régression beta-binominale a été appliqué pour analyser l'effet du groupe (CFL vs Copilot) sur la proportion de réponses correctes (fonction `glmmTMB` dans R). Les prédictions du modèle ont ensuite permis d'estimer la proportion moyenne de réponses correctes pour chaque groupe (voir Figure 1 ci-dessous): 56 % pour le groupe Copilot et 51 % pour le groupe CFL. Un second modèle de régression a été estimé afin d'examiner l'effet du groupe sur le nombre total de réponses produites, à l'aide de la fonction `glmmTMB`. Le modèle suggère que le groupe CFL a suggéré moins d'items, mais cette différence n'était pas statistiquement significative (valeur de $P = 0,21$).

Nous avons ensuite entré tous les mots récoltés par les apprenant·es dans l'outil d'analyse de complexité lexicale de FLELex (François et al. 2014; Pintard & François 2020)⁸ pour en tirer un profil selon les niveaux du CECR: niveau-cible A1-B1, B2, C1 et C2. Nous avons comparé ces profils de niveaux à celui des verbes proposés par les dictionnaires utilisés dans cette étude (voir Fig. 2). Nous avons utilisé un test du chi-carré pour comparer les distributions des niveaux pour chaque groupe (CFL, Copilot, Dictionnaire). Ces tests ont montré que ces différences étaient statistiquement significatives (CFL vs Dictionnaire, valeur de $P = < 2.2e-16$; Copilot vs Dictionnaire, valeur de $P = < 2.2e-16$).

⁸ URL: <https://cental.uclouvain.be/cefrlex/flelex/analyse/>

3.3 Résultats

3.3.1 Verbes récoltés par les apprenant·es

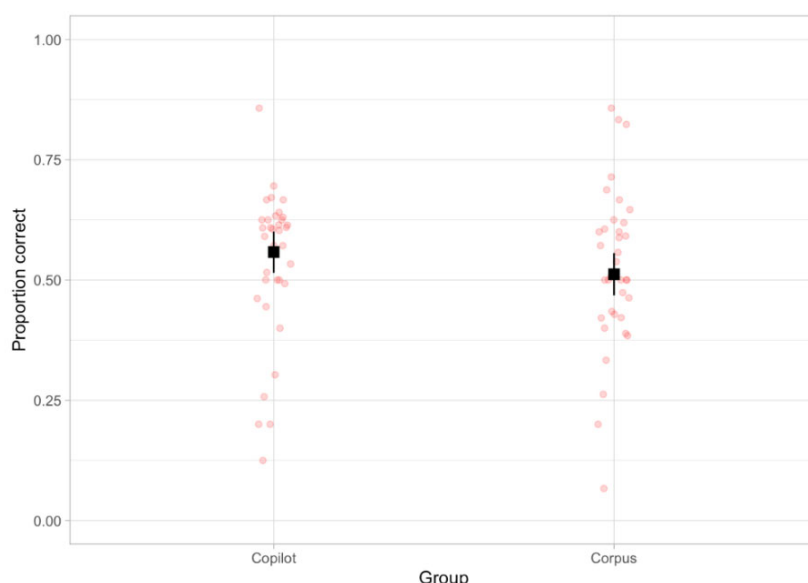


Figure 1: Visualisation du modèle prédictif (carré noir), avec plus ou moins d'incertitude (flèches noires), créé à partir des données brutes (en gris) pour le nombre de réponses correctes pour chaque groupe

Nous nous sommes d'abord penchées sur le nombre de réponses correctes dans les deux groupes. Le groupe Copilot a en moyenne un taux de réponses correctes plus élevé, mais cette différence n'est pas statistiquement pertinente (valeur de $P = 0.14$) et dans le modèle prédictif les intervalles de confiance du modèle pour les deux groupes se chevauchent. Il en va de même pour le nombre de mots récoltés par les deux groupes. Les différences sont très petites et ne sont pas statistiquement pertinentes.

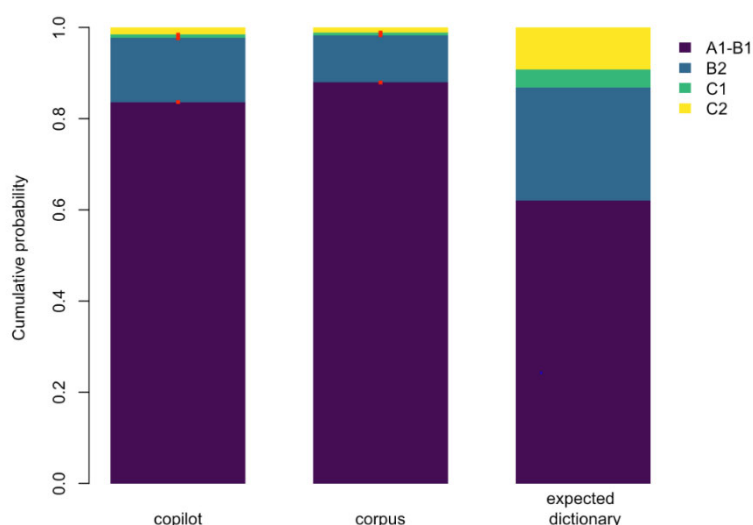


Figure 2: Les proportions de verbes proposés dans chaque catégorie de niveau pour le groupe CFL, le groupe Copilot et pour les dictionnaires utilisés.

Comme on le voit dans la figure 2, le nombre de verbes de niveau B2 proposé par les apprenant·es du groupe Copilot est un peu plus important que pour le groupe CFL. Mais ce qui frappe est surtout la différence entre les niveaux des verbes rassemblés par les étudiant·es et ceux des dictionnaires. Ces différences sont statistiquement significatives (voir-ci dessus) et montrent que les apprenant·es proposent plus de mots de niveaux A1-B1 que les dictionnaires.

3.3.2 Les réponses au questionnaire

Pour toutes les questions, un modèle de régression logistique a été appliqué afin d'évaluer si le groupe (CFL vs Copilot) exerçait une influence significative sur les résultats. Le modèle a été ajusté avec la fonction glm en utilisant une distribution binomiale. Malgré de légères différences constatées, les résultats statistiques ont montré que les différences entre les réponses des deux groupes aux questions à choix multiples n'étaient pas statistiquement significatives, (voir figure 3).

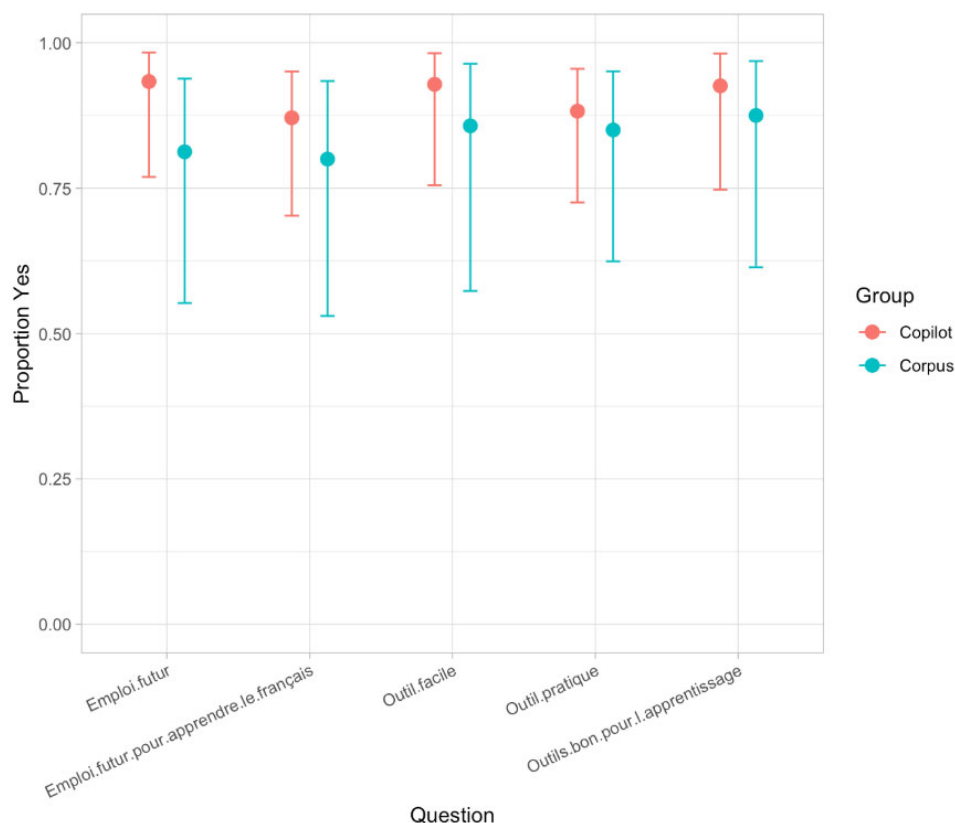


Figure 3. Résultats du modèle prédictif pour la réponse oui aux cinq questions principales.

Pour chaque question, les réponses ont été globalement très positives, comme le montre la figure 4. Malgré ces réponses positives, un membre du groupe CFL critique l'outil dans sa réponse à la question ouverte, estimant: "Je pense que je n'utiliserai pas cet outil à l'avenir, il ne m'a pas plu. Aujourd'hui, il existe des outils plus pratiques et faciles à utiliser."

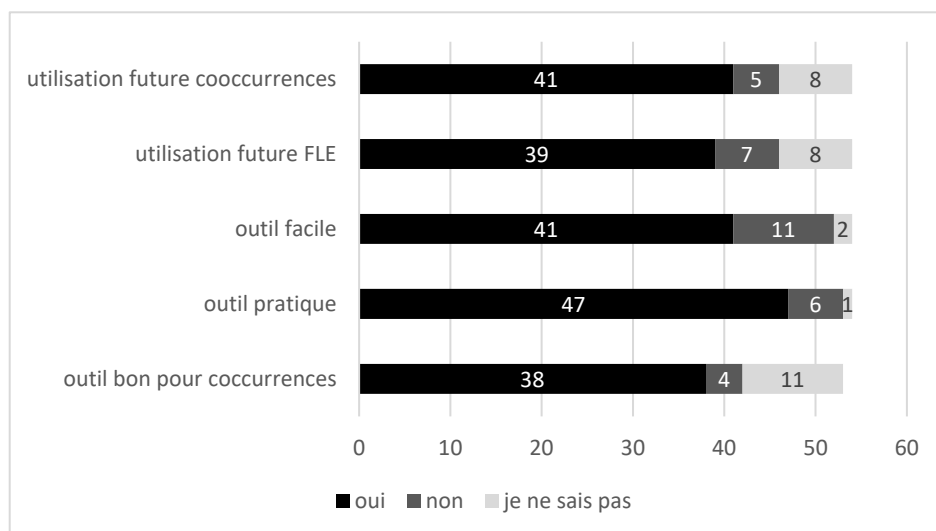


Figure 4: Résultats des réponses au questionnaire (n=54).

La question ouverte s'intéressait à la façon dont les apprenant·es allaient utiliser l'outil (CFL ou Copilot) dans le futur. Seules 34 personnes ont répondu à cette question et les réponses des deux groupes sont assez diverses. Pour le groupe Copilot, elles reflètent les différentes utilisations possibles des GAT, notamment "écrire des textes" (Et8), faire des comparaisons avec la langue maternelle (Et14)⁹, "corriger" des textes (Et5, Et19), créer des exemples (Et64)¹⁰, utilisation pour des tâches qu'ils ne peuvent faire eux·elles-mêmes (Et73) ou pour clarifier des informations (Et48)¹¹. Le groupe CFL a partagé moins de réponses (11/34) dont trois qui mentionnent qu'ils ne l'utiliseront pas et deux qu'ils n'avaient pas bien compris la tâche. Les autres réponses montrent surtout une utilisation pour l'apprentissage du lexique (synonyme, collocation, vocabulaire¹²).

⁹ Original: "J'observe des variantes de combinaisons de mots en français et en russe. Cela rend le processus d'apprentissage très facile pour moi."

¹⁰ Original: "Pour chercher les constructions verbales, pour demander, crée exemple pour mieux comprendre."

¹¹ Original: "utilisation occasionnelle pour absorber des informations et clarifier quelque chose de peu clair."

¹² Originaux: "Je vais l'utiliser cet outil pour communiquer et compléter votre vocabulaire de mots", "En cherchant des synonymes", "Je vais utiliser cet outil pour vérifier et apprendre les collocations avec les mots différents."

4. Discussion

Les résultats non statistiquement pertinents entre les deux groupes sur de nombreux points tendent à montrer que l'utilisation de l'IAGen n'est pas plus efficace que le travail avec les corpus pour la recherche de cooccurrences par des apprenant·es de niveau B1. Le nombre de mots recueillis est un peu plus élevé mais la différence n'est pas statistiquement pertinente et il en est de même pour le nombre de mots corrects. Lorsqu'on se penche dans le détail des verbes collectés, on remarque toutefois une différence significative. À l'exception d'un·e étudiant·e du groupe Copilot qui n'a manifestement pas compris comment rédiger un prompt et obtient de la part du GAT des résultats problématiques (voir Fig. 5), les autres membres du groupe mentionnent uniquement des verbes. Ce n'est pas le cas pour les propositions du groupe CFL. Par exemple, pour le mot *métier*, un nombre assez important de membres

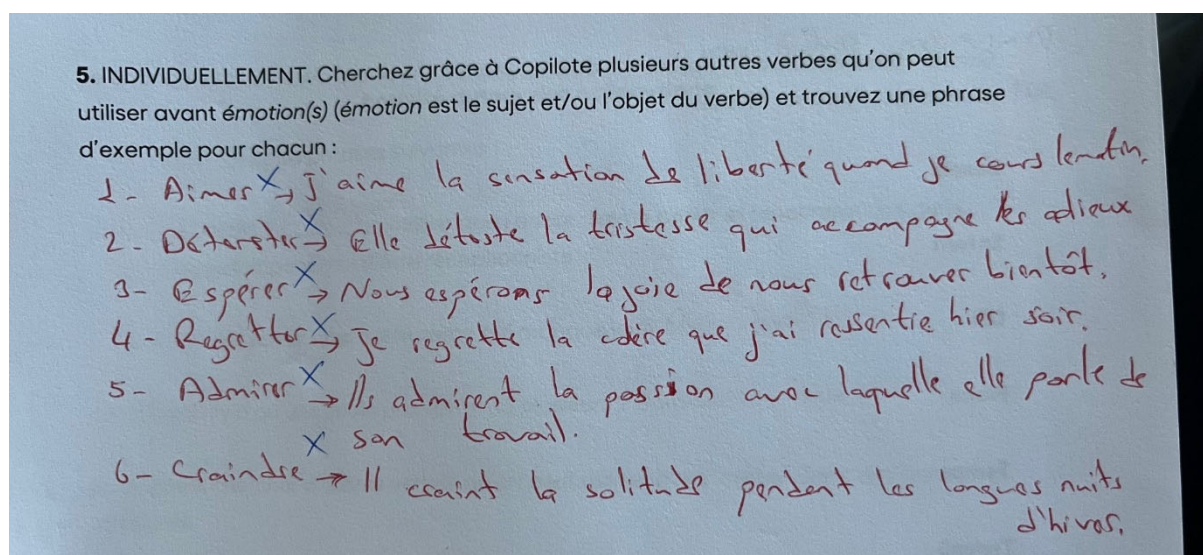


Figure 5: Reproduction de l'exercice où un·e étudiant·e du groupe Copilot devait noter des verbes qui entraient en cooccurrence avec *émotion(s)* et noter une phrase d'exemple. Vu le résultat retranscrit par l'étudiant·e, le prompt utilisé devait s'éloigner du modèle donné en classe et ne devait pas être efficace, car aucune des phrases proposées – qui sont par ailleurs des phrases correctes – ne comprend le mot *émotion*, même si elles décrivent des émotions.

du groupe CFL mentionne deux items inattendus: *tisser* (13 occurrences) et *savoir-faire* (11 occurrences). Il semble probable que les étudiant·es n'aient pas compris le mot composé *métier à tisser* et ait interprété celui-ci comme *métier* + verbe. Pour *savoir-faire*, on peut imaginer que cet autre nom composé de deux verbes apparaisse dans les voisins de droite ou de gauches significatifs en raison de son sémantisme. Un autre problème récurrent avec le groupe CFL est l'utilisation du même verbe avec une flexion différente. Ainsi, Et33 propose pour le mot *besoin* les verbes *ai*, *ont*, *avez*, *aura*, *as* avec chaque fois une phrase d'exemple tout à fait pertinente, mais sans se rendre compte qu'il s'agissait du même verbe. Ces quelques exemples d'erreurs de 'casting' dans

le groupe CFL pourraient nous faire penser que l'utilisation d'un GAT est plus pertinente pour cet exercice. Bien évidemment, lorsque la tâche est accompagnée par un·e enseignant·e, ces erreurs permettent des discussions intéressantes avec les apprenant·es.

Néanmoins, les résultats obtenus sur le profil des niveaux des verbes recueillis par les participant·es pointent plutôt vers le corpus. En effet, le groupe CFL a collecté plus de verbes dans le niveau-cibles (c.à.d. entre A1 et B1) et moins de verbes d'un niveau plus élevé qui ne sont pas pertinents pour des apprenant·es de niveau B1. Il est vrai que les prompts qui ont été partagés ne mentionnaient pas forcément le niveau attendu (même si nous avons insisté sur l'importance d'un contexte spécifique dans la courte formation en prompting donnée aux participant·es et indiqué le niveau dans le prompt d'examen). Peut-être qu'un meilleur prompt qui indique le niveau (débutant-moyen inférieur) aurait pu donner des résultats plus utiles pour les apprenant·es? Il conviendrait maintenant de mener une recherche plus ciblée sur la compréhension ou non par les GAT des niveaux du CECR, que de nombreux·euses enseignant·es utilisateurs·trices de ces outils considèrent comme allant de soi.

Un résultat intéressant est la comparaison des deux outils, CFL et GAT, avec les dictionnaires de cooccurrences utilisés. On voit que les livres de référence ont tendance à présenter des verbes qui sont d'un niveau plus élevé que ce que proposent les outils basés sur les algorithmes statistiques. Ainsi, au niveau B1, il est plus productif pour l'apprenant·e de consulter un corpus ou un GAT plutôt que de se tourner vers un dictionnaire. À nouveau, le résultat n'est statistiquement significatif que pour les deux dictionnaires utilisés. Il serait intéressant de tenter l'expérience avec d'autres ouvrages de référence, comme le *Dictionnaire de cooccurrences* de *Druide* informatique par l'équipe d'Antidote dans sa version digitale et imprimée (Charest et al. 2012), ou le *Dictionnaire combinatoire compact du français* (Zinglé & Brobeck-Zinglé 2011).

Les réponses au questionnaire ne montrent pas de différences statistiquement significatives entre les deux groupes, même si le nombre de réponses positives est un peu plus élevé dans le groupe Copilot. Cette légère préférence pour le GAT se lit également dans les réponses à la question ouverte où plusieurs membres du groupe CFL expriment leur désintérêt, voire leur antipathie pour cet outil, en écho aux problèmes évoqués par Boulton (2009, 2017).

5. Conclusion

Le parcours théorique et expérimental que nous avons proposé dans cet article suggère un bilan plutôt mitigé pour l'intégration du produit de l'IAGen pour l'enseignement/apprentissage des LE. On l'a vu dans la première partie, les problèmes qui persistent dans le produit des GAT et sa révision, et qui sont inhérents à la technologie de ces outils, contrebalancent les gains de temps

extraordinaires souvent vantés par les utilisateurs·trices. Bien sûr, plus de recherches doivent maintenant être entreprises sur différents modèles de GAT pour (in)valider ces premières observations.

Pour la partie expérimentale, on voit que, si le GAT paraît pertinent et facile d'utilisation aux apprenant·es, il ne l'est pas beaucoup plus que les corpus, un outil qui a déjà démontré sa pertinence à entrer en classe de LE. Pour certain·es participant·es qui ne savent pas prompter, le GAT n'a permis de produire aucune réponse correcte. Ce n'est jamais le cas pour le corpus. Même si quelques participant·es plus récalcitrant·es semblent préférer l'IAGen, le corpus recueille de nombreuses impressions positives. Les résultats des étudiant·es des deux groupes sont similaires et les GAT ne sont pas meilleurs. Le nombre similaire d'items recueillis dans les deux groupes tend aussi à montrer que les GAT ne permettent pas aux apprenant·es d'être plus efficaces et de gagner du temps dans la recherche de cooccurrences. À l'inverse, les GAT proposent plus de termes qui sont d'un niveau trop élevé pour nos participant·es B1.

Ces résultats suggèrent que l'intégration des IA en classe nécessite un encadrement pédagogique précis qui comprend, notamment le prompt engineering et le développement pour les apprenant·es de la littéracie en IAGen (Cotelli Kureth et al. 2025). On sait que l'introduction de nouvelles technologies dans l'apprentissage peut créer un certain stress pour les étudiant·es mais des études suggèrent le rôle central de l'enseignant·e pour atténuer le technostress des étudiant·es (Yang et al. 2025).

Notre enquête permet ainsi de casser la frénésie et les annonces fracassantes qui ont parfois accompagné l'introduction de l'IAGen dans l'enseignement/apprentissage des LE. Les GAT sont, en effet, utiles pour certaines tâches, comme la recherche des cooccurrences, mais elles ne sont pas tellement plus faciles et plus rapides d'utilisation que d'autres outils qui ont déjà montré leur efficacité dans ces tâches. Nous craignons en effet parfois que le côté 'couteau-suisse' des GAT qui peuvent tout faire (rédiger, corriger, traduire, reformuler, synthétiser, résumer, informer, etc.) ne fassent oublier les nombreux outils spécialisés qui existent déjà et ont prouvé leur efficacité. Il convient finalement de lire ces résultats mitigés en parallèle aux risques éthiques et environnementaux que nous avons mentionnés et nous demander pour toute utilisation des GAT, même si elle est possible et/ou utile, si elle est toujours nécessaire et pertinente.

Pour conclure, il est donc crucial de récolter plus de données pour garantir une utilisation pertinente de l'IAGen dans l'enseignement/apprentissage des LE. En parallèle, les enseignant·es de LE et les apprenant·es doivent impérativement développer une littéracie en IAGen qui comprend, selon le modèle de Cardon et collègues (2023) l'authenticité et la responsabilité, mais aussi l'application qui permet aux utilisateurs·trices de l'IA de connaître quel outil fonctionne le mieux

pour quelle tâche et l'agentivité qui leur permet de garder le contrôle de l'interaction avec la machine et son produit.

6. Déclarations et remerciements

Nous n'avons utilisé aucune IAGen pour la recherche de données, l'analyse et la rédaction de cet article, à l'exception d'Antidote pour une révision grammaticale et orthographique et Copilot pour nous aider dans la médiation en français du rapport statistique.

Nous remercions le Linguistic Research Infrastructure (LiRI)¹³ de l'Université de Zürich) qui a effectué les analyses et les calculs statistiques sur nos données. Nous remercions également nos étudiant·es qui ont accepté de partager le résultat de leurs recherches et leurs avis sur les outils. Cette recherche a été possible grâce au soutien financier du projet swissuniversities "Digital Literacy in University Contexts" (2021-2024)¹⁴.

BIBLIOGRAPHIE

- Adams, C., Pente, P., Lernermeier, G. & Rockwell, G. (2023). Ethical principles for artificial intelligence in K-12 education. *Computers and Education: Artificial Intelligence*, 4, 100131. <https://doi.org/10.1016/j.caeai.2023.100131>
- Ancil, D. & Tremblay, O. (2016). Les collocations: des combinaisons de mots privilégiées. *Correspondance*, 21(3), sans pagination. <https://correspo.ccdmd.qc.ca/wp-content/uploads/2018/09/correspondance-la-lecture-dans-tous-ses-etats-les-collocations-des-combinaisons-de-mots-privilegiees-.pdf>
- Benites, F., Delorme Benites, A. & Anson, C. M. (2023). Automated text generation and summarization for academic writing. In O. Kruse, C. Rapp, C. Anson et al. (éds.), *Digital writing technologies in higher education. Theory, research, and practice* (pp. 279-301). Cham: Springer.
- Boulton, A. (2007). Esprit de corpus: Promouvoir l'exploitation de corpus en apprentissage des langues. *Texte et corpus*, 3, 37-46. hal-00326988
- Boulton, A. (2009). Documents authentiques, oral, corpus. *Mélanges CRAPEL*, 31, 5-13.
- Boulton, A. (2017). Data-driven learning and language pedagogy. In S. Thorne & S. May. (éd.) *Language, education and technology: Encyclopedia of language and education*, 3, 181-192. Cham: Springer
- Boulton, A., Canut, E., Guerin, E., Parisse, Ch. & Tyne, H. (2015). Corpus et appropriation de L1 et L2. *LINX*, 68/69, 9-32.
- Boulton, A. & Tyne, H. (2014). *Des documents authentiques aux corpus: démarches pour l'apprentissage des langues*. Paris: Didier.
- Brown, H., Tarone, E., Swan, M., Ellis, R., Jung, U., Prodromou, L., Bruton, A., Johnson, K., Nunan, D., Oxford, R., Goh, C., Waters, A. & Savignon, S. (2007). Forty years of language teaching. *Language Teaching*, 40, 1-15.

¹³ URL: <https://www.liri.uzh.ch/en.html>

¹⁴ URL: <https://www.zhaw.ch/en/linguistics/research/digital-literacy-in-university-contexts>



- Cardon, P., Fleischmann, C., Aritz, J., Logemann, M. & Heidewald, J. (2023). The challenges and opportunities of AI-assisted writing: Developing AI literacy for the AI age. *Business and Professional Communication Quarterly*, 86(3). 235-405.
- Cavalla, C. (2019). Une entrée lexicale spécifique: les collocations. In M. L. Perassi, M. Zapia Kwiecien (éds.). *Palabras como puentes: Estudios lexicológicos, lexicográficos, y terminológicos desde el Cono Sur* (pp. 58-67). Córdoba: Buena Vista Editora.
- Charest, S., Fontaine, J. & Saint-Germain, J. (2012). *Le grand druide des cooccurrences. Dictionnaire*. Montréal: Éditions Druide.
- Cotelli Kureth, S., Delorme Benites, A., Lehr, C., Nogherechi, H., Summers, E. & Steele, E. (2025). Comment développer la littéracie digitale des enseignant-es et des apprenant-es? Conclusions du projet 'Digital Literacy in University Contexts'. *Babylonia*, 56-59. <https://doi.org/10.55393/babylonia.v1i.510>
- Cotelli Kureth, S. & Nogherechi, H. (2023). How inclusivity can help develop machine translation (MT) literacy and autonomy: A Swiss experiment with learners of French. *Instructed Second Language Acquisition*, 7(2), 268-280.
- Cotelli Kureth, S. & Nogherechi, H. (2024). Intégrer GoogleTranslate comme support d'enseignement au niveau débutant: atelier pratique pour un public de réfugié·e·s ukrainien·ne·s. *Partages*, 1, en ligne. <https://dx.doi.org/10.35562/partages.98>
- Crompton, H. & Burke, D. (2023). Artificial intelligence in higher education: the state of the field. *International Journal of Educational Technology in Higher Education*, 20(1), 1-22.
- François, T., Gala, N., Watrin, P. & Fairon, C. (2014). FLELex: A graded resource for French foreign learners. In *Proceedings of the 9th International Conference on Language Resources and Evaluation (LREC 2014)*. Reykjavik, Iceland, 26-31 May, paper 1108. http://www.lrec-conf.org/proceedings/lrec2014/pdf/1108_Paper.pdf
- Gilquin, G. & Granger, C. (2010). How can data-driven learning be used in language teaching? In A. O'Keeffe & M. McCarthy (éds.), *Routledge handbook of corpus linguistics* (pp. 359-370). London: Routledge.
- Gilquin, G. & Gries, S. (2009). Corpora and experimental methods: A state-of-the-art review. *Corpus Linguistics and Linguistic Theory*, 5(1), 1-26.
- Gouvernement du Québec (2024). *L'utilisation pédagogique, éthique et légale de l'intelligence artificielle générative. Guide destiné au personnel enseignant*. Québec: Ministère de l'Éducation. URL: <https://cdn-contenu.quebec.ca/cdn-contenu/education/Numerique/Guide-utilisation-pedagogique-ethique-legale-IA-personnel-enseignant.pdf>
- Hovy, D., Bianchi, F. & Fornaciari, T. (2020). "You sound just like my father": Commercial machine translation systems include stylistic biases. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 1686-1690). Association for Computational Linguistics. <https://aclanthology.org/2020.acl-main.154.pdf>
- Hulstijn, J.H. (2013). Incidental learning in second language acquisition. In C.A. Chapelle (éd.), *The encyclopedia of applied linguistics* (v.5, pp. 2632-2640). Weinheim: Wiley-Blackwell. <https://doi.org/10.1002/9781405198431.wbeal0530>
- Ji, M. (2023). Cultural and linguistic bias of neural machine translation technology. In M. Ji, P. Bouillon & M. Seligman (éd.), *Translation technology in accessible health communication* (pp. 100-128). Cambridge: Cambridge University Press.
- Johns, T. (1997). Kibbitzing one-to-ones. Web version. University of Reading: BALEAP on Academic Writing. <http://www.eisu.bham.ac.uk/johnstf/pimnotes.htm>.
- Johns, T. & King, Ph. (éds.) (1991). Classroom concordancing. *English Language Research Journal*, 4, 27-45.

- Jolley, J. & Maimone, L. (2022). Thirty years of machine translation in language teaching and learning: a review of the literature. *L2 Journal*, 14(1), 26-44.
- Kamber, A. (2011). Contexte et sens: utilisation d'un corpus écrit dans l'enseignement/apprentissage du FLE. *Travaux Neuchâtelois de Linguistique*, 55, 199-218.
- Klimova, B., Pikhart, M., Delorme Benites, A., Lehr, C. & Sanchez-Stockhammer, C. (2023). Neural machine translation in foreign language teaching and learning: A systematic review. *Education and Information Technologies*, 28, 663-682. <https://doi.org/10.1007/s10639-022-11194-2>
- Kurokawa, S. & Salingre, M. (2025). Syntactic and lexical comparison between AI-generated reading passages and Japanese universities' national test. *Journal of Studies in Language, Culture, and Society (JSLCS)*, 8(1); 295-309.
- Lee, S.-M. & Briggs, N. (2021). Effects of using machine translation to mediate the revision process of Korean university students' academic writing. *ReCALL*, 33(1), 18-33. <https://doi.org/10.1017/S0958344020000191>
- Leray, M. & Tyne, H. (2016). Homophonie et maîtrise du français écrit: apport de l'apprentissage sur corpus. *Linguistik online*, 78(4), 131-148. <https://www.redalyc.org/articulo.oa?id=664573544009>
- Look, R. (2020). No more rage against the machine: How corpus-based identification of machine-translationese can lead to student empowerment. *The Journal of Specialised Translation*, 34, 150-170. https://www.jostrans.org/issue34/art_look.php
- Look, R. & Léchauguet, S. (2021). Machine translation literacy and undergraduate students in applied languages: Report on an exploratory study. *Revista Tradumàtica: tecnologies de la traducció*, 19, 204-225. <https://doi.org/10.5565/rev/tradumatica.281>
- Galindo-Dominguez, H., Delgado, N., Losada, D. & Etxabe, J. (2023). An analysis of the use of artificial intelligence in education in Spain: The in-service teacher's perspective. *Journal of Digital Learning in Teacher Education*, 40(1), 41-56. <https://doi.org/10.1080/21532974.2023.2284726>
- Holt, B., Look, R. & Léchauguet, S. (2022). The use of online translators by students not enrolled in a professional translation program: Beyond copying and pasting for a professional use. In *Proceedings of the 23rd Annual Conference of the European Association for Machine Translation* (pp. 23-29). Ghent: European Association for Machine Translation. <https://aclanthology.org/2022.eamt-1.5>
- Mair, C. (2002). Empowering non-native speakers: the hidden surplus value of corpora in continental English departments. In B. Kettemann & G. Marko (éds.), *Teaching and learning by doing corpus analysis* (pp. 119-130), Amsterdam & New York: Rodopi.
- Martindale, M. J. & Carpuat, M. (2018). Fluency over adequacy: A pilot study in measuring user trust in imperfect MT. In *Association for Machine Translation in the Americas 2018*. <https://doi.org/10.48550/arXiv.1802.06041>
- Molle, N. & Hamza-Jamann, A. (2025). Cartographie des usages de l'Intelligence Artificielle Générative (IAG) dans l'apprentissage des langues chez les étudiants Lansad. *Recherches en didactique des langues et des Cultures*, 23(2). <http://journals.openedition.org/rdlc/16438>; <https://doi.org/10.4000/15bf1>
- Motlagh, N. Y., Khajavi, M., Sharifi, A. & Ahmadi, M. (2023). The impact of artificial intelligence on the evolution of digital education: A comparative study of OpenAI text generation tools including ChatGPT, Bing Chat, Bard, and Ernie. arXiv preprint. <https://arxiv.org/abs/2309.02029>
- Nesselhauf, N. (2004). Learner corpora and their potential for language teaching. In J. Sinclair (éd.), *How to use corpora in language teaching* (pp. 125-152), Amsterdam & Philadelphia: Benjamins.
- O'Sullivan, Í. (2007). Enhancing a process-oriented approach to literacy and language learning: the role of corpus consultation literacy. *ReCALL*, 19(3), 269-86.
- Pintard, A. & François, T. (2020). Combining expert knowledge with frequency information to infer CEFR levels for words. In *Proceedings of the 1st workshop on tools and resources to empower people*

- with *REAding Difficulties (READI)* (pp. 85-92). Marseille: European Language Resources Association. <https://aclanthology.org/2020.readi-1.13/>
- Primerano, A. (2022). L'émergence des concepts de "capacitisme" et de "validisme" dans l'espace francophone. *ALTER*, 16(2), 43-58. <https://journals.openedition.org/alterjdr/683?file=1>
- Raaflaub, M. & Reber, B. (2025). Machine translation literacy in the lower secondary classroom. *Bulletin suisse de linguistique appliquée*, 120, 65-78.
- Richards, K. (2009). Trends in qualitative research in language teaching since 2000. *Language Teaching*, 42(2), 147-180.
- Sarrion, E. (2023). *Exploring the power of ChatGPT: Applications, techniques and implications*. New York: Apress/Springer. <https://doi.org/10.1007/978-1-4842-9529-8>
- Savoldi, B., Gaido, M., Bentivogli, L., Negri, M. & Turchi, M. (2021). Gender biases in machine translation. *Transactions of the Association for Computational Linguistics*, 9, 845-874. https://doi.org/10.1162/tac1_a_00401
- Shaitarova, A. (2023). Generated or human: Towards coherence assessment in German and English news articles. Presented at the 3d LITHME conference, at the University of Groningen in Leeuwarden, on May 15, 2023.
- Shi, R., Grissom, A. II & Trinh, D. (2022). Rare but severe neural machine translation errors induced by minimal deletion: An empirical study on Chinese and English. In *Proceedings of the 29th international Conference on Computational Linguistics* (pp. 5175-5180). Gyeongju: International Committee on Computational Linguistics. <https://aclanthology.org/2022.coling-1.459>
- Sun, W. & Park, E. (2023). EFL learners' collocation acquisition and learning in corpus-based instruction: A systematic review. *Sustainability*, 15(17), 13242. <https://doi.org/10.3390/su151713242>
- Thomas, A., Granfeldt, J., Jouin-Chardon, E. & Etienne, C. (2016). Conversations authentiques et CECR: compréhension globale d'interactions naturelles par des apprenants de FLE. *Cahiers de l'AFLS*, 20(1). <https://du.diva-portal.org/smash/record.jsf?pid=diva2%3A913133&dswid=5036>
- Tsai, S.-C. (2019). Using google translate in EFL drafts: A preliminary investigation. *Computer Assisted Language Learning*, 32(5-6), 510-526. <https://doi.org/10.1080/09588221.2018.1527361>
- Tyne, H. (2013). Corpus et apprentissage-enseignant des langues. *Bulletin suisse de linguistique appliquée*, 97, 7-15.
- Wang, Y. & Daghigh A. J. (2023). Cognitive effort in human translations and machine translation post-editing processes. A holistic and phased view. *FORUM. Revue internationale d'interprétation et de traduction*, 21(1), 139-162.
- Yang, M., Wu, X. & Deris F. D. (2025). Exploring EFL learners' positive emotions, technostress and psychological well-being in AI-assisted language instruction with/without teacher support in Malaysia. *British Educational Research Journal*, <https://doi.org/10.1002/berj.4184>
- Zinglé, H. & Brobeck-Zinglé, M.-L. (2011). *Dictionnaire combinatoire compact du français*. Paris: La maison du dictionnaire.