Revue TDFLE

N° 85 | 2025
La gestion de l'apport culturel par les technologies en didactique du FLE : enjeux, défis et perspectives

L'apport culturel dans la reconnaissance automatique de la parole en didactique du FLE

Alida Maria SILLETTI *Maître de conférences*
Département de Sciences politiques
University of Bari Aldo Moro

Édition électronique :
URL :
https://revue-tdfle.fr/articles/revue-85/452-l-apport-culturel-dans-la-reconnaissance-automatique-de-la-parole-en-didactique-du-fle
DOI : numerev_2573
Date de publication : 04/07/2025

Cette publication est sous licence **CC BY-NC-ND** (Attribution - No commercial - No derivatives).

Pour **citer cette publication** : SILLETTI, A. M. (2025) L'apport culturel dans la reconnaissance automatique de la parole en didactique du FLE. *Revue TDFLE*, (85). https://doi.org/10.34745/numerev_2573

Cette recherche relève de l'emploi d'outils issus de l'intelligence artificielle (IA) à des fins didactiques en classe de FLE. Elle consiste à vérifier la manière dont le logiciel de transcription automatique de YouTube traite de l'apport culturel à partir de huit messages-vidéo du Président de la République française, prononcés entre 2023 et 2024 et disponibles sur le site de l'Élysée. La caractérisation de ce corpus, relevant du discours institutionnel, est effectuée à l'appui de l'analyse du discours de l'école française. Ces messages-vidéo présentent une charge culturelle importante qui est manifestée tant par les thèmes abordés, dont le but est surtout de commémorer des événements d'envergure de l'histoire contemporaine française ou européenne, que par la manière dont Emmanuel Macron s'y réfère en termes langagiers. L'analyse est notamment conduite à partir d'une sélection de « culturèmes » et de mots isolés, d'expressions ou de passages prononcés dans une langue différant du français. Le traitement de l'apport culturel de la part du logiciel d'IA de YouTube s'y avère être déficitaire pour diverses raisons, mettant en cause aussi bien l'entraînement du logiciel de transcription automatique que son fonctionnement. Celui-ci est biaisé par la non-prise en compte du contexte et par des facteurs qui sont liés aux variabilités affectant le signal acoustique.

Abstract

This research deals with the use of artificial intelligence (AI) for didactical purposes as far as French as a foreign language is concerned. It consists of analysing the way in which YouTube's speech-to-text tool performs in identifying the cultural dimension which emerges from eight video-messages by the French President Emmanuel Macron, pronounced between 2023 and 2024 and published on the Presidency's website. These messages belong to institutional discourse and are examined through the methodology of French discourse analysis. They contain a strong cultural dimension, depending not only on their topic – whose aim is mainly to commemorate recent French and European historical events – but also on the way in which E. Macron refers to them. The analysis will particularly look at some "culturemes" and isolated words, expressions or sentences which are not pronounced in French. It will be shown that the cultural dimension of these messages is not taken into account by YouTube AI tool for several reasons, among which the functioning of this tool and its biases, consisting of its indifference towards the context and of all the variabilities related to the acoustic signal.

Keywords

Speech-to-text tools ; cultural dimension ; FLE ; institutional discourse ; artificial intelligence

Mots-clés :
FLE, Intelligence artificielle, Discours institutionnel, Transcription automatique, Apport culturel

Introduction

La présente étude s'inscrit dans une recherche pilote que nous conduisons depuis l'année universitaire 2020-2021, en Italie, en didactique du français comme langue étrangère (FLE), au niveau universitaire (master), filières LANSAD, par le biais de et pour promouvoir l'exploitation du numérique dans la didactique des langues étrangères. Elle est encadrée dans des situations d'enseignement « innovantes » (Lameul et Loisy, 2014) en didactique, relevant des potentialités de l'intelligence artificielle (IA) en termes de reconnaissance automatique de la parole continue. Si l'essor de l'IA dans tous les domaines des activités humaines et de la recherche est incontournable, dans le domaine linguistique ses applications majeures ont trait à l'industrie des langues. Le but est de viser des objectifs différents en se servant d'approches statistiques d'apprentissage profond fondées sur des corpus (Chaumartin et Lemberger, 2020 ; Raus et Tonti, 2025), parmi lesquels celui de favoriser l'apprentissage des langues étrangères via la traduction et la transcription automatiques.
Si les travaux sur l'utilisation de l'IA à des fins d'apprentissage se multiplient, ils s'intéressent de plus en plus également au choix des données d'entraînement des algorithmes qui sont à la base des outils d'IA. L'apprentissage de la machine s'avère être partiel en raison de plusieurs facteurs, parmi lesquels l'impossibilité d'interpréter des textes et donc de tenir compte du contexte, ainsi qu'un fonctionnement qui est basé sur des codes (Rastier, 2021). D'où l'enjeu du type d'apprentissage de la machine et l'importance de l'intervention humaine en amont, afin que cet apprentissage soit supervisé et que ses résultats soient moins biaisés. Cela met en cause pour autant également la part de responsabilité de l'être humain en aval des résultats qui sont issus de l'outil d'IA, c'est-à-dire lors du processus de post-édition. C'est ce dernier processus qui est davantage exploité en classe de langue.
Cette recherche s'intéressera à l'ensemble de ces réflexions, en les rapportant à un petit corpus que nous exploitons lors de l'atelier de transcription automatique qui est proposé à nos élèves. Bien que cette activité pédagogique vise des objectifs premièrement langagiers au niveau de la morphosyntaxe, de la syntaxe de la phrase complexe et de la prosodie de la langue française (Silletti, 2023 ; 2025 ; soumis) à partir des problèmes qui émergent des transcriptions automatiques de l'outil de reconnaissance automatique de YouTube, elle présente également un intérêt vis-à-vis de l'apport culturel. Aux fins de cette recherche, nous entendons par « apport culturel » des éléments exprimés par la langue, autrement dit des passages, des expressions ou des mots, même isolés, dont la charge culturelle est importante. Ceux-ci sont inscrits dans des événements discursifs (Maingueneau, 2021) d'envergure et sont donc

culturellement (et idéologiquement) connotés. Nous aurons notamment affaire à la communication institutionnelle du Président de la République française via huit messages-vidéos prononcés entre 2023 et 2024 et disponibles sur le site de l'Élysée. Ce genre de discours est caractérisé non seulement par sa brièveté et par une « scène d'énonciation » (Maingueneau, 2021) stable, comme nous le montrerons au long de l'analyse, mais aussi par des propos pourvus d'une charge culturelle importante. Ces messages-vidéos portent en effet surtout sur des anniversaires et sur des commémorations d'événements d'envergure qui ont marqué l'histoire nationale, européenne ou supranationale, ou bien sur des salutations officielles qu'E. Macron adresse lors d'événements internationaux auxquels il participe à distance depuis l'Élysée. Nous nous intéresserons à la manière dont ce matériel, qui n'est pas toujours présenté en langue française, est appréhendé par le logiciel d'IA de YouTube en termes de résultats des transcriptions automatiques. Cela permettra également de réfléchir sur la qualité des données issues de ce logiciel d'IA et sur le choix des données qui ont été utilisées en amont pour entraîner l'algorithme. De surcroît, pour qu'un biais ou une erreur soient identifiés, il est indispensable de s'attarder sur ce qui est considéré comme biais ou erreur et sur leur catégorisation, afin de mieux évaluer les résultats obtenus.

Notre méthodologie s'appuiera donc, d'une part, sur la relation entre technologies numériques et enseignement/apprentissage des langues étrangères ; d'autre part, sur la communication institutionnelle du Président de la République et sur le genre de discours du message-vidéo à partir de l'analyse du discours de l'école française. Quant aux caractéristiques de l'outil de transcription automatique de YouTube, après avoir montré le fonctionnement des logiciels contemporains de reconnaissance automatique de la parole en continu, nous illustrerons les points de force et les faiblesses de ce système (Tancoigne *et al*., 2020 ; Silletti, 2025), qui se répercutent inexorablement sur les résultats des transcriptions automatiques, tous niveaux confondus.

1. L'apport du numérique dans l'apprentissage des langues étrangères

L'appui sur des outils d'intelligence artificielle à des fins didactiques permet de viser de compétences qui sont souvent négligées en classe de langue, parmi lesquelles la prise en compte de l'apport culturel.

Face aux positionnements sur les changements induits par le numérique dans l'enseignement et l'apprentissage des langues, rappelés par Develotte (2022), loin de sous-estimer l'enseignement traditionnel des langues étrangères, notre démarche se situe à la croisée entre une approche qui permet de faire converger l'« innovation didactique » et l'« innovation technologique » vers une « innovation durable » (Puren, 2022, en ligne). Si, en effet, par le modèle du « déterminisme didactique », qui voit le jour dès le début du XXe siècle, ce sont les besoins de chaque discipline qui amènent les spécialistes qui les enseignent à se servir de l'outillage technologique en tant que « moyens auxiliaires » pour réaliser les objectifs attendus, par le modèle du « déterminisme technologique » (Puren, 2022, en ligne), introduit depuis la fin des années 1950 à partir de l'enseignement de l'anglais comme langue étrangère, c'est

l'intégration entre technologies différentes qui engendre des pratiques didactiques innovantes auprès du corps enseignant. Le web contribue désormais à la démocratisation de ces technologies, en les rendant disponibles partout et pour tout le monde, à condition de posséder une connexion en ligne, préconisant même l'idée d'après laquelle l'intégration de l'intelligence artificielle aux outils technologiques permettrait à la machine de remplacer l'enseignement humain. C'est de l'intégration en termes de convergence entre ces deux modèles qu'il sera question dans la présente étude. Nous partageons le point de vue de Puren (2022), qui propose une diffusion large, voire pérennisée, de l'innovation technologique dans les pratiques enseignantes qui « converge » avec une innovation didactique, mais en continuant à attribuer un rôle essentiel à l'enseignement humain, qui reste incontournable.

Notre expérience didactique pilote d'exploitation d'outils numériques issus de l'IA à des fins didactiques voit le jour à partir d'un double constat : l'importance de l'innovation dans les pratiques langagières d'apprentissage du FLE et la diffusion de plus en plus capillaire de l'IA dans les pratiques étudiantes de la part d'un public qui n'est pas toujours formé à faire un emploi raisonné des outils issus de l'IA. C'est ce qui est entre autres émergé à partir de la soumission de questionnaires sur l'utilisation d'outils de traduction et de transcription automatiques à nos élèves, dans le cadre des activités du centre d'excellence Jean Monnet *Artificial Intelligence for European Integration* (http://www.jmcoe.unito.it/home). Les résultats de ces questionnaires, élaborés dans le cadre de ce projet de recherche, vont d'une confiance relative en ces outils, par rapport à laquelle l'intervention humaine est considérée comme fondamentale, à l'indistinction entre dictionnaires en ligne et traducteurs automatiques (Raus et Silletti, 2022) en aval de l'exploitation de ces outils en classe de FLE. D'où notre volonté d'intégrer des outils de l'IA à nos pratiques enseignantes, en vue de faire converger innovation technologique et innovation didactique, mais sans que la première ait le dessus sur la seconde. Pour ce faire, nous considérons également indispensable l'apport des corpus en didactique des langues (Auzéau et Abiad, 2018).

2. Un corpus pourvu d'une charge culturelle élevée

L'apport des corpus dans les processus d'apprentissage est désormais reconnu et la « réaction instinctive » de méfiance que Boulton (2007) constatait à cet égard, en 2007, par rapport à la langue française, est contrecarrée par leur exploitation tant dans l'enseignement des langues que dans la formation des élèves pour apprendre une langue (Fligelstone, 1993).

Il est évident que le choix du corpus est directement proportionnel à l'analyse à conduire et à la possibilité d'y retrouver des régularités en vue de confirmer ou de rejeter des hypothèses préalables. Dans le cadre de la présente recherche, nous avons cherché à identifier, par rapport aux genres de discours de la communication politico-institutionnelle du Président de la République française, ceux dont la charge culturelle est élevée, dans le but d'en vérifier le traitement par l'outil d'IA de YouTube. Nous avons constaté que la présence de références culturelles est plus élevée dans des messages-vidéos que dans des discours officiels, dans des conférences de presse ou dans des interviews du Président de la République. Les huit messages-vidéo que nous

avons collectés à partir du portail de l'Élysée (www.elysee.fr), datant d'août 2023 à mai 2024, concernent en effet des événements commémoratifs, des anniversaires, et, plus rarement, la participation du Président de la République à des initiatives nationales, européennes ou internationales, par rapport auxquelles il présente les salutations officielles de la Nation. Cette parole présidentielle, qui est en partie préparée et qui sera étudiée via sa transcription automatique, s'inscrit, comme nous l'avons constaté dans Silletti (soumis), dans une « scène d'énonciation » (Maingueneau, 2021) qui témoigne d'un discours « unifié », « lissé » (Oger et Ollivier-Yaniv, 2006) et en partie planifié. Sa « scène englobante » (Maingueneau, 2021), à savoir le cadre pragmatique minimal qui précise les conditions d'énonciation inscrites dans le genre de discours concerné, relève à la fois du discours politique et de la communication institutionnelle d'un sujet politique jouant un rôle institutionnel bien identifié. Ce sujet est le seul qui est légitimé à prendre la parole, qu'il ne doit pas négocier avec un sujet interlocuteur en face de lui. Par ailleurs, ces messages-vidéos sont enregistrés, ce qui empêche toute forme directe et immédiate de confrontation. Ils sont inscrits dans une « scénographie » (Maingueneau, 2021) spécifique, caractérisé par un même lieu d'énonciation, à savoir le Palais de l'Élysée, et par un même arrière-plan faisant supposer qu'il s'agit du bureau du Président de la République. Celui-ci est en effet pourvu de symboles institutionnels et « identitaires », parmi lesquels les drapeaux français et européen (et éventuellement celui du pays concerné par l'événement commémoré), et le sceau de la République française. Le sujet énonciateur montre également une « posture présidentielle » : il apparaît au premier plan, assis ou debout, en costume, regardant une caméra fixe, sans bruits de fond et par rapport à un temps de parole de cinq minutes en moyenne par message. Les destinataires de cette élocution peuvent correspondre, selon les circonstances d'énonciation, aux responsables politiques d'autres pays, à des organismes et institutions nationaux ou supranationaux, ainsi qu'à la population française ou à une partie de celle-ci.
Avant de nous attarder sur la dimension culturelle qui est le propre des messages-vidéos consultés, nous présenterons le fonctionnement des logiciels de reconnaissance automatique de la parole en continu, en vue de mieux comprendre l'outil de transcription automatique de YouTube.

3. La reconnaissance automatique de la parole en continu

Le domaine de la reconnaissance automatique de la parole a subi des évolutions remarquables grâce au développement des technologies et aux performances de plus en plus accrues des ordinateurs (Heba, 2021). Les premiers systèmes de traitement automatique de la parole, qui portent sur la reconnaissance de mots isolés, remontant aux années 1960, sont caractérisés par une composante acoustique, phonétique et linguistique sur base statistique. Ces systèmes ne sont pourtant pas en mesure de traiter de la différence de voix, de la vitesse d'élocution, de l'accent, voire de la prononciation du même mot de la part de sujets différents en raison de caractéristiques sociales, physiologiques ou environnementales – correspondant à la « variabilité inter-locuteur » (Barrault, 2008) –, mais aussi des incohérences liées à la langue parlée. Pour y faire face, c'est d'abord le traitement de phrases qui est proposé, en utilisant un

vocabulaire plus important et des unités phonétiques en nombre limité, pour ensuite aller vers un modèle statistique incluant non seulement les modèles acoustiques statistiques mais également les grammaires statistiques. C'est donc par un calcul probabiliste que le système est tenu de reproduire une phrase, un mot ou des segments sonores et qu'une modélisation statistique avancée peut voir le jour (Heba, 2021).

Depuis les années 2010, l'introduction de l'« apprentissage profond », permettant à un dispositif basé sur des réseaux neuronaux qui imitent ceux du cerveau humain d'apprendre via des données (Le Cun, 2019), comporte la possibilité de prédire des sorties textuelles directement à partir d'entrées audio, sans devoir construire des composants intermédiaires relevant des modèles distincts pour les sons élémentaires (partie acoustique), les séquences de mots (partie linguistique) et les prononciations (lien entre la partie acoustique et linguistique), cherchant à simuler des performances humaines. Ses avantages principaux relèvent de l'épargne de temps et d'un traitement qui concerne également des mots hors du vocabulaire, trouvant la meilleure transcription pour un segment audio donné (Heba, 2021).

Par conséquent, l'intégration de modules fondés sur les réseaux neuronaux artificiels a permis d'accroître les performances des systèmes de reconnaissance automatique de la parole (Tancoigne *et al*., 2020). Pourtant, un taux d'imperfection y est toujours présent. Celui-ci est fortement lié aux déformations ou variations, ayant des effets divers sur le signal de parole, et affectant la suite de mots la plus probable, c'est-à-dire le résultat produit par le logiciel utilisé (Barrault, 2008). Elles dépendent de l'environnement dans lequel l'enregistrement a lieu, de l'état et du mode d'expression du sujet locuteur lui-même et des conditions d'enregistrement (Haton, 2006), mais aussi de la taille et de la difficulté du vocabulaire. D'autres problèmes concernent le traitement automatique de la langue naturelle courante, en raison de la nécessité de prendre en compte le contexte (Mariani, 1990). Par ailleurs, les transcriptions automatiques nécessitent un long travail humain de préparation et d'entraînement des logiciels, ce qui pose la question de l'entraînement des données et du choix de celles-ci, mais aussi d'une vérification humaine en amont et en aval de la conception de l'automate pour son bon fonctionnement.

Relativement aux systèmes actuels de transcription automatique pour la langue française, Tancoigne *et al*. (2020) comparent huit logiciels, dont celui de YouTube, basé sur la technologie Google Voice. Depuis 2009, YouTube diffuse des sous-titres automatiques qui naissent, au départ, pour aider les personnes sourdes et malentendantes. Dans ce système, les sous-titres sont calculés à l'avance en prenant en compte l'intégralité de la bande acoustique. Ce système est entraîné à partir de millions d'heures d'enregistrement de données audio et avec des milliards de phrases de texte, générant des transcriptions automatiques activables sur demande et gratuites. Celles-ci apparaissent sous forme de sous-titres automatiques en bas de l'écran ou bien comme texte projeté à côté de l'écran, d'où la possibilité de visualiser la transcription au fur et à mesure que l'énonciation se poursuit. En effet, celle-ci est accompagnée des balises temporelles et peut faire l'objet d'un copier-coller. L'analyse conduite par Tancoigne *et al*. (2020), portant entre autres sur l'identification, le classement et la correction des erreurs relevées dans les transcriptions automatiques des logiciels examinés, à laquelle nous ajoutons nos travaux préalables (Silletti, 2023 ;

2025 ; soumis) sur le logiciel de YouTube, soulignent que ces erreurs ont trait à la substitution d'un mot par rapport à ce qui est énoncé, à la suppression et donc à l'omission d'un mot, mais aussi à l'ajout de mots non prononcés. C'est ce même classement des erreurs, qui repose d'abord sur la perception auditive, qui sera appliqué à la présente étude, notamment à des aspects lexico-sémantiques. Il est utile de préciser que Tancoigne *et al*. (2020) ne traitent pas de l'apport culturel affectant les logiciels de transcription automatique examinés.

Dans notre atelier, afin d'effectuer l'exercice de révision des transcriptions automatiques, après avoir visualisé et écouté la vidéo du message à examiner, les élèves réfléchissent sur la transcription automatique à partir d'un fichier word – que nous leur préparons au préalable – contenant un tableau avec deux colonnes. Celle de gauche présente la transcription automatique, tandis que celle de droite doit être remplie par la transcription révisée effectuée par chaque élève en post-édition. L'exercice qui leur est demandé est ainsi comparable à une traduction intralinguistique en parallèle. C'est sous ce même format que seront présentés les cas qui font l'objet de notre analyse (Fig. 1 et suivantes).

4. Analyse du corpus : un apport culturel mal maîtrisé

Parmi les huit messages-vidéo du Président de la République, quatre portent sur des événements commémoratifs, trois relèvent d'événements internationaux par rapport auxquels le Président de la République présente les salutations officielles de la France à ses homologues et un seul message, dont la portée est plus nationale, concerne ses propos lors de la Journée internationale de lutte contre les violences faites aux femmes. Ces prises de parole témoignent donc d'un potentiel culturel important, dont il faudra vérifier le traitement par le logiciel d'IA de YouTube. À cet effet, deux cas problématiques seront présentés : les « culturèmes » et les formulations qui sont le résultat de l'appui sur des langues-cultures autres que le français.

4.1 Les « culturèmes »

Les premiers cas de transcription automatique erronée qui ont suscité notre intérêt concernent la non-identification, de la part du logiciel d'IA, d'unités linguistiques porteuses d'informations culturellement marquées, qui naissent dans une langue-culture pour être traduites littéralement, par calque, dans d'autres langues-cultures. Nous nous référons à des « culturèmes » (Pamies, 2017, entre autres), que nous concevons dans une définition étendue aux fins de la présente traduction intralinguistique entre oral prononcé et oral transcrit. Le logiciel d'IA de YouTube tend à ne pas reconnaître ces unités, les remplaçant par des mots lexicalisés dans la langue-culture française et donc ne renvoyant pas aux référents culturels auxquels elles font référence. Le résultat de la transcription automatique correspond à un non-sens sémantique dû à une erreur par substitution de mots, que seule la connaissance de cette référence culturelle permettra de corriger en post-édition. C'est ce qui apparaît dans le message-vidéo de commémoration pour le 50e anniversaire de la Révolution des Œillets, renvoyant à un moment historique d'envergure dans l'histoire européenne de l'après-Seconde guerre mondiale : le renversement de la dictature de Antonio

Salazar, au Portugal, et le retour à la démocratie, le 25 avril 1974. Emmanuel Macron rappelle cet événement 50 ans après, en le désignant par le syntagme nominal « Révolution des Œillets », qu'il prononce au début de son message-vidéo. Cette occurrence n'est pourtant pas correctement saisie par le logiciel de transcription automatique de YouTube (Fig. 1) : les fleurs symbole de cette révolution, les œillets, sont remplacées par un adjectif homophone, « œillés » – accordé au pluriel en raison de la liaison obligatoire qui est prononcée par le locuteur –, générant un non-sens sémantique et donc une erreur par substitution.

Figure 1 (Non-reconnaissance des « culturèmes » par le logiciel d'IA de YouTube)[1] *(https://www.elysee.fr/emmanuel-macron/2024/04/24/50eme-anniversaire-de-la-revolution-des-oeillets-le-message-du-president-emmanuel-macron)*[2]

Transcription automatique	Transcription révisée
0:00 demain marquera les 50 ans de la 0:02 **Révolution des œillés** et je souhaite me 0:05 joindre à nos amis du Portugal au 0:06 Portugais de France et à tous nos […]	0:00 Demain marquera les 50 ans de la 0:02 Révolution des Œillets. Et je souhaite me 0:05 joindre à nos amis du Portugal, aux 0:06 Portugais de France et à tous nos […]

Il en va de même, dans ce message-vidéo, pour le renvoi aux « Capitaines d'avril », autrement dit de jeunes officiers qui, avec la population civile, ont libéré leur pays de la dictature salazariste – « Capitaines d'avril » correspond également au titre d'un long-métrage franco-portugais réalisé en 1999, renvoyant à ce même événement. Or, quoique les mots composant le syntagme aient été bien saisis par le logiciel d'IA, c'est le culturème qui fait défaut par le manque de l'initiale en majuscules de « Capitaines » dans les deux occurrences de « Capitaines d'avril » (dont nous ne reproduisons que la première, Fig. 2) figurant dans le même message-vidéo. Le résultat, c'est une erreur par substitution, bref un autre non-sens sémantique :

Figure 2 (Non-reconnaissance des « culturèmes » par le logiciel d'IA de YouTube) (https://www.elysee.fr/emmanuel-macron/2024/04/24/50eme-anniversaire-de-la-revolution-des-oeillets-le-message-du-president-emmanuel-macron)

Transcription automatique	Transcription révisée
3:23 nous le devons à la mémoire des 3:24 **capitaines d'avril** et au peuple 3:26 portugais qui il y a 50 ans s'est 3:28 soulevé contre l'injustice et la 3:30 et a suscité l'admiration de la France 3:32 de l'Europe et du monde un rouge à la 3:36 main avril VI Portugal au brigade au 3:41 Portugal	3:23 Nous le devons à la mémoire des 3:24 Capitaines d'avril et au peuple 3:26 portugais qui, il y a 50 ans, s'est 3:28 soulevé contre l'injustice et la pauvreté, 3:30 et a suscité l'admiration de la France, 3:32 de l'Europe et du monde, un œillet rouge à la main. 25 de abril, sempre ! Viva Portugal ! Obrigado a Portugal !

Malgré la diffusion de ces deux syntagmes représentant autant de culturèmes de la langue-culture portugaise même au-delà des frontières nationales et dans la langue-culture française, le logiciel de YouTube a produit d'une part un homophone (« œillés »), d'autre part le syntagme mais sans le traiter en tant que culturème (par le manque de l'initiale majuscule). Cela nous fait supposer que les bases de données sur lesquelles est entraîné le logiciel de YouTube (relevant de Google) ne contiennent pas ces référents culturels ou en tout cas un nombre d'attestations qui suffise à leur reconnaissance et saisie correcte de la part de la machine.
Ces cas ne sont pas isolés dans les messages-vidéos examinés et se vérifient même lorsque le nom de l'événement à commémorer porte sur un culturème qui est directement lié à la France et qui présente une dénomination en français. « Entente cordiale », qui fait référence à l'accord politique scellé le 8 avril 1904 par la France et par le Royaume-Uni pour régler les divergences coloniales entre les deux puissances, dont le 8 avril 2024 E. Macron commémore, par un message-vidéo, le 120ème anniversaire,

Figure 3 (Non-reconnaissance de « culturèmes » par le logiciel d'IA de YouTube) (https://www.elysee.fr/emmanuel-macron/2024/04/08/releve-franco-britannique-de-la-garde-a-loccasion-des-120-ans-de-lentente-cordiale)

Transcription automatique	Transcription révisée
0:00 il y a 120 ans aujourd'hui le 8 avril 0:02 1904 le Royaume-Uni et la France 0:04 concluaient **l'entente cordiale** balayant 0:08 les rivalités coloniales d'alors et les 0:09 siècles de relations tumultueuses du 0:11 […]	0:00 Il y a 120 ans aujourd'hui, le 8 avril 0:02 1904, le Royaume-Uni et la France 0:04 concluaient l'Entente cordiale, balayant 0:08 les rivalités coloniales d'alors et les 0:09 siècles de relations tumultueuses du 0:11 […]

est correctement saisi par le logiciel d'intelligence artificielle à chacune de ses trois occurrences, il est bien transcrit en termes grammaticaux, mais il n'est jamais reconnu en tant qu'unité culturellement connotée. En effet, son initiale n'est pas en majuscules, et c'est toute sa charge historique, culturelle, voire émotionnelle qui est perdue. Or, tant cet exemple que le précédent témoignent d'une utilisation non systématique des signes graphiques au sein du logiciel d'IA de YouTube, lesquels apparaissent par exemple, en termes d'initiales en majuscules, à propos de « France » et « Royaume-Uni » (Fig. 3). Les attestations de ces noms propres de pays sont sans aucun doute déjà présentes dans le logiciel d'IA de YouTube, leur fréquence est plus importante et leur prononciation respecte la langue française standard, d'où leur saisie correcte par le logiciel de transcription automatique.
La non-prise en compte de l'apport culturel de la part du logiciel d'IA de YouTube se manifeste également *via* le recours à des mots, à des expressions ou à des passages issus d'autres langues-cultures dans les messages-vidéos du Président de la République.

4.2 Les formulations venant d'autres langues-cultures

Au-delà des cas que nous venons de présenter, qui prouvent d'abord un manque d'attestations (suffisantes), dans les données d'entraînement du logiciel d'IA de YouTube, d'occurrences de culturèmes portant sur l'histoire même nationale française, l'apport culturel des messages-vidéo examinés peut être identifié dans deux autres situations. Celles-ci partagent l'appui sur des référents exprimés par des codes linguistiques différant du français.
La première concerne des dénominations exprimées dans une langue distincte du français, qui ne sont pas saisies (erreurs par omission) ou qui sont mal transcrites (erreurs par substitution) par le logiciel de transcription automatique de YouTube. Tel est le cas d'éléments isolés, à savoir des noms propres de personne, des toponymes, des noms d'associations, des noms de chansons.
« MeToo/Me Too », à savoir le nom du mouvement féministe qui, depuis octobre 2017, se bat contre les violences faites aux femmes, qui est né aux États-Unis et s'est ensuite répandu dans le monde entier, est cité par le Président de la République lors de son

message-vidéo publié à l'occasion de la Journée internationale de lutte contre les violences faites aux femmes, le 25 novembre 2023. Bien que le nom du mouvement soit correctement prononcé par E. Macron, par ailleurs précédé de son apposition « mouvement », cet emprunt n'est pas saisi par le logiciel d'IA de YouTube, qui génère une erreur par substitution. « MeToo » est remplacé par « Mitou » (Fig. 4), à savoir un homophone respectant les règles de prononciation du français, pourvu d'une initiale en majuscules et donc pouvant se référer à un nom propre. D'après une recherche sur le web, « Mitou » pourrait à la limite correspondre à un nom propre de personne – il en est ainsi de Maître Mitou, célèbre organiste vécu au XVIe siècle – ou de lieu – c'est, par exemple, l'une des traductions du mont Mitoku japonais. Le résultat est un autre non-sens sémantique.

Figure 4 (Non-identification de référents culturels par homonymie) (https://www.elysee.fr/emmanuel-macron/2023/11/25/journee-internationale-de-lutte-contre-les-violences-faites-aux-femmes-le-message-du-president-emmanuel-macron)

Transcription automatique	Transcription révisée
0:34 toutes les formes de violence le 0:36 mouvement **Mitou** est venu libérer la 0:38 parole des femmes des victimes ont osé […]	0:34 toutes les formes de violence. Le 0:36 mouvement MeToo est venu libérer la 0:38 parole des femmes. Des victimes ont osé […]

Nous excluons que la non-identification de « MeToo » dans cet exemple découle de la manière dont le Président de la République prononce cette formulation en anglais, étant donné qu'elle est homophone par rapport au résultat proposé par l'outil d'IA de YouTube. Nous attribuons plutôt ce biais de l'algorithme aux données d'entraînement de celui-ci, qui présentent sans doute un nombre d'occurrences de « MeToo » trop bas pour que cette formulation soit la solution la plus probable, même si elle apparaît au sein d'un syntagme pourvu également de son apposition, contribuant davantage à lever l'ambiguïté. Il serait pour autant également utile de s'interroger sur « Mitou », renvoyant à des référents qui sont beaucoup moins répandus que « MeToo ».
Relativement aux noms propres de personne et de lieu, ils sont rarement bien saisis par le logiciel d'IA de YouTube, même s'ils puisent dans l'histoire et dans la culture françaises récentes ou contemporaines.
Cependant, les problèmes liés à l'onomastique et à la toponymie ne sont souvent pas dus au fait que ces formulations proviennent de langues différant du français. Dans le message-vidéo sur la commémoration du 80ème anniversaire de la Résistance en France, en 2023, des noms et des lieux symboles de cette période de l'histoire de France ne sont pas transcrits correctement par le logiciel d'IA de YouTube. Tel est le cas du militaire et maquisard français Jean Moulin, que le logiciel d'IA de YouTube rend par un homonyme, « jeanmoulin », sans séparation entre le prénom et le nom et faute d'initiales en majuscules (Fig. 5) :

Figure 5 (Non-identification de référents culturels par homonymie) (https://www.elysee.fr/emmanuel-macron/2024/03/06/80eme-anniversaire-de-la-liberation-le-message-du-president-emmanuel-macron)

Transcription automatique	Transcription révisée
0:00 nous avons commémoré tout au long de 0:01 l'année 2023 la Résistance française au 0:04 plus fort de l'oppression avec les 0:05 hommages rendus à **jeanmoulin** mais aussi 0:07 aux **enfants 10 yeux** à la prison de 0:10 Montluc à toutes les victimes plus 0:12 largement de la barbarie donc à Lyon 0:14 dans cette prison au Conseil national de 0:16 la Résistance **rue du fourour** à Paris aux 0:19 résistants de Corse tout 0:21 particulièrement à **ajaacio** aux étrangers […]	0:00 Nous avons commémoré tout au long de 0:01 l'année 2023 la Résistance française au 0:04 plus fort de l'oppression, avec les 0:05 hommages rendus à Jean Moulin, mais aussi 0:07 aux enfants d'Izieu, à la prison de 0:10 Montluc, à toutes les victimes plus 0:12 largement de la barbarie, donc, à Lyon 0:14 dans cette prison, au Conseil national de 0:16 la Résistance, rue du Four à Paris, aux 0:19 résistants de Corse, tout 0:21 particulièrement à Ajaccio, aux étrangers […]

Un autre cas d'homonymie correspondant à un autre non-sens sémantique touche au culturème pourvu d'un toponyme « enfants d'Izieu », renvoyant à la rafle de la colonie d'Izieu et à l'arrestation, puis à la déportation, de 44 enfants et sept adultes juifs par la Gestapo. La transcription automatique du logiciel de YouTube produit un homonyme, générant une erreur par substitution, en raison de la non-identification du culturème et de la reproduction d'une formulation française existante, mais hors contexte. Quant à « rue du Four » et « Ajaccio », qui figurent dans le même passage mais que le logiciel de transcription automatique ne saisit pas correctement, les erreurs par substitution générées par le système d'IA sont le résultat d'une « variabilité intra-locuteur ». Ce type de variation, que nous définissons à partir de ce que Barrault (2008) qualifie de « variabilité inter-locuteur », concerne le fait que la prononciation de la même expression par le même sujet locuteur sera toujours différente en raison de la prosodie, de l'intonation, de l'articulation, mais aussi du cotexte où elle apparaît.

De tels résultats sont encore plus fréquents si les événements commémorés, et l'onomastique et la toponomastique qui y sont associées, concernent l'histoire et la culture d'autres pays et les formulations dont il est question sont exprimées dans la langue de ces pays. Dans le message-vidéo commémorant la Révolution des Œillets,

c’est ce qui arrive lorsqu’E. Macron rappelle, par leurs noms, des femmes et des hommes héroïques dans ce combat pour la liberté (Fig. 6) :

Figure 6 (Identification erronée de noms étrangers de personne par le logiciel d’IA de YouTube)
(https://www.elysee.fr/emmanuel-macron/2024/04/24/50eme-anniversaire-de-la-revolution-des-oeillets-le-message-du-president-emmanuel-macron)

Transcription automatique	Transcription révisée
1:41 lui avoir offert l'asile je pense 1:44 aujourd'hui à toutes ces femmes à tous 1:47 ces hommes qui ont partagé sa cause je 1:50 pense aux trois Maria **Maria Isabelle** 1:52 **barenau Maria Théresa Horta** et **Maria véo** 1:55 **d'Acosta** dont le combat pour les droits 1:57 des femmes portugaises sous les 1:59 dictatures a reçu un soutien immense 2:04 en France et a en quelque sorte annoncé 2:07 le 25 avril je pense à **Mario Suarez** 2:11 **Alvaro cugal** et **Mido Guerrero** qui ont 2:16 rejoint Lisbonne et Porto dans les jours 2:17 qui ont suivi la révolution à partir de […]	1:41 lui avoir offert l'asile. Je pense 1:44 aujourd'hui à toutes ces femmes, à tous 1:47 ces hommes qui ont partagé sa cause ; je 1:50 pense aux trois Maria : Maria Isabel 1:52 Bareno, Maria Teresa Horta et Maria Velho 1:55 da Costa, dont le combat pour les droits 1:57 des femmes portugaises sous les 1:59 dictatures a reçu un soutien immense 2:04 en France, et a, en quelque sorte, annoncé 2:07 le 25 avril. Je pense à Mário Soares, 2:11 Álvaro Cunhal et Emidio Guerreiro, qui ont 2:16 rejoint Lisbonne et Porto dans les jours 2:17 qui ont suivi la révolution, à partir de […]

Ces erreurs par substitution, qui ne relèvent pas toujours d’homonymes, peuvent dépendre de la non-prise en compte du contexte et de la probabilité d’une absence de ces noms propres pourvus d’une forte charge culturelle dans le logiciel d’IA de YouTube. En outre, la non-reproduction correcte par le logiciel de ce qui est énoncé peut résulter de la manière dont le sujet locuteur prononce ces noms. Ces formulations sont en portugais et, hormis les prénoms qui pourraient être également répandus en France, tels que « Mario/Maria » et « Alvaro », par rapport aux noms de famille le logiciel a cherché le plus possible à adapter leur transcription à la langue française. Ainsi, il n’est pas possible d’exclure que si ces noms avaient été prononcés par un sujet lusophone leur transcription aurait pu être correcte. Le logiciel d’IA de YouTube est entraîné à

partir d'une prononciation, d'une intonation et d'une prosodie basées sur la variante standard/nationale de la langue concernée.
Cela peut être confirmé par les derniers cas de formulations dotées d'une charge culturelle importante, en raison de la langue d'expression utilisée et de ce qui est sous-tendu par ce choix en termes tant culturels et diplomatiques que de transcription automatique. E. Macron se fait le porte-parole d'un discours de commémoration qui est demandé par les conditions qui sont inscrites dans la production du message-vidéo, qu'il a tendance à rapprocher du pays concerné et de l'événement commémoré. À cela contribuent des expressions prononcées dans la langue relative à « l'autre », dans une volonté de partage mutuel. Ce positionnement permet au Président de la République de mieux sceller les relations diplomatiques entre les pays concernés dans le cas des commémorations et lors des événements internationaux auxquels la France participe à distance. Ainsi, lors du Forum mondial de l'Eau, à Bali, le 21 mai 2024, E. Macron envoie un message-vidéo de 5,39 minutes au Président indonésien Joko Wi. À la fin de son message, après avoir souhaité un bon déroulement des travaux du forum, le Président français remercie pour l'attention en indonésien, par « terima kasih », à savoir « merci ». Cette séquence verbale est investie d'enjeux sociopolitiques et fonctionne en tant que terme de relation, mais elle n'est pas reconnue par le logiciel de transcription automatique, qui ni ne la saisit ni ne la transcrit correctement (Fig. 7). Le non-sens qui en résulte investit également « à Bali », transcrit par un homonyme – alors que, dans les autres occurrences figurant dans ce même message, « Bali » est correctement transcrit :

Figure 7 (Non-identification de séquences verbales isolées en langue étrangère par le logiciel d'IA de YouTube) (https://www.elysee.fr/emmanuel-macron/2024/05/21/forum-mondial-de-leau-de-bali)

Transcription automatique	Transcription révisée
5:30 détermination de la France pour faire de 5:33 l'eau une question majeure au cœur de 5:35 l'agenda international alors bon forum à 5:39 tous et à toute abali **trimakassi**	5:30 détermination de la France pour faire de 5:33 l'eau une question majeure au cœur de 5:35 l'agenda international. Alors, bon forum à 5:39 tous et à toutes à Bali. Terima kasih.

Cette erreur par substitution dépend aussi bien des données d'entraînement du logiciel d'IA de YouTube que de la prononciation de cette formulation en indonésien par le sujet locuteur, qui n'est pas censé maîtriser de cette langue.
La plupart des messages-vidéos examinés contiennent des expressions relevant d'une langue autre que le français, correspondant à celle du pays concerné par les événements à commémorer ou à évoquer. Ainsi, E. Macron conclut son message-vidéo pour le 50e anniversaire de la Révolution des Œillets en exhortant le Portugal à poursuivre son combat pour la démocratie et en le remerciant, en portugais. *Via* cette formulation, il montre la reconnaissance de la France à l'égard de la lutte pour la liberté

du peuple portugais. Encore une fois, le logiciel d'IA de YouTube produit un non-sens résultant d'erreurs par substitution et par omission, là où cet outil a cherché à reproduire en langue française ce qui a été prononcé par le sujet locuteur (Fig. 8). Pour autant, nous estimons que la confusion générée par le logiciel d'IA de YouTube est surtout due à la variabilité inter-locuteur, c'est-à-dire à une prononciation imparfaite de cette expression portugaise par E. Macron, et donc à sa saisie et transcription automatiques incorrectes.

Figure 8 (Non-identification de séquences verbales isolées en langue étrangère par le logiciel d'IA de YouTube) (https://www.elysee.fr/emmanuel-macron/2024/04/24/50eme-anniversaire-de-la-revolution-des-oeillets-le-message-du-president-emmanuel-macron)

Transcription automatique	Transcription révisée
3:24 capitaines d'avril et au peuple 3:26 portugais qui il y a 50 ans s'est 3:28 soulevé contre l'injustice et la 3:30 et a suscité l'admiration de la France 3:32 de l'Europe et du monde un rouge à la 3:36 main **avril VI Portugal au brigade au** 3:41 Portugal	3:24 Capitaines d'avril et au peuple 3:26 portugais qui, il y a 50 ans, s'est 3:28 soulevé contre l'injustice et la pauvreté, 3:30 et a suscité l'admiration de la France, 3:32 de l'Europe et du monde, un œillet rouge à la main. 25 de abril, sempre ! Viva Portugal ! Obrigado a Portugal !

À cela, il est possible d'ajouter une autre variable qui influence les résultats du logiciel d'IA de YouTube, à savoir la durée, très brève, de l'élocution dans une langue autre que le français, prononcée au sein d'une suite acoustique en français.
À notre avis, ces biais ne sont pas influencés par le choix de la langue étrangère utilisée per le sujet locuteur, bien qu'il maîtrise sans aucun doute mieux l'anglais que le portugais et l'indonésien. Cela se vérifie même si langue étrangère concernée est l'anglais, à savoir la principale langue d'entraînement des logiciels d'IA, toutes applications confondues (Raus et Tonti, 2025, entre autres), dont la transcription automatique est incorrecte s'il n'y a pas de correspondance entre le signal acoustique du logiciel d'IA et celui qui est perçu à partir de l'oral à transcrire. Pour en témoigner, il suffit de vérifier la transcription automatique du passage en anglais, long de presque une minute (dont nous ne reproduisons que le début, Fig. 9), qu'E. Macron prononce à la fin du message-vidéo célébrant le 120e anniversaire de l'Entente cordiale. Son but est de donner une image « participative » de son pays et de soi pour montrer le lien d'amitié entre les deux pays et tout ce que recèle la langue-culture. Pour autant, le logiciel d'IA de YouTube, qui ne prend en compte ni le contexte ni ce qui est sous-tendu par cette alternance de langues, produit des non-sens. D'entiers passages sont omis,

tandis que les autres sont caractérisés par des erreurs par substitution :

Figure 9 (Non-reconnaissance de séquences verbales isolées en anglais par le logiciel d'IA de YouTube) (https://www.elysee.fr/emmanuel-macron/2024/04/08/releve-franco-britannique-de-la-garde-a-loccasion-des-120-ans-de-lentente-cordiale)

Transcription automatique	Transcription révisée
2:28 juillet prochain dans le cadre de la 2:30 communauté politique européenne que nos 2:33 amis britanniques vont 2:36 organiser let's 2:42 hope 2:45 togetherers of normand prov landing 2:48 liberation of 2:51 France his of 2:58 sacrifice […]	2:28 juillet prochain, dans le cadre de la 2:30 communauté politique européenne que nos 2:33 amis britanniques vont 2:36 organiser. Let's 2:42 hope it shines this summer on the other anniversary we shall be celebrating 2:45 together this year: the 80th anniversary on the Normandy and Provence landings, 2:48 and the liberation of 2:51 France. We will never forget our shared history of 2:58 sacrifice and comembership. We should […]

Or, au-delà d'une transcription entièrement à réviser en post-édition, l'écoute de ce passage montre que ces résultats sont en partie dus à la variabilité inter- et aussi intra-locuteur : le même mot ne sera jamais prononcé de la même manière par le même sujet locuteur et le logiciel d'IA est entraîné à partir de la variante standard/nationale de la langue concernée. Au fond, il est également possible de confirmer que, comme Tancoigne *et al*. (2020) le relèvent, l'outil de transcription automatique de YouTube cherche à transcrire tout ce qu'il perçoit, en dépit du sens de l'original (le pourcentage de reproduction avoisine 80-85 % de ce qui est prononcé).

Il s'ensuit que ce n'est pas la présence de plusieurs codes linguistiques à la fois qui pose des problèmes pour le logiciel d'IA de YouTube, mais la prononciation, l'intonation et la prosodie de ces langues – et sans doute la durée de cette énonciation. Il émerge aussi que les données d'entraînement des outils issus de l'IA ne couvrent pas l'intégralité des domaines d'emploi d'une langue. L'histoire contemporaine, tant nationale qu'européenne et internationale, y est ainsi peu représentée au profit de sujets qui portent, entre autres, sur l'IA elle-même, dont les attestations sont plus nombreuses. En témoigne, dans le même corpus de messages-vidéos, la saisie correcte du sigle « IA », y compris ses initiales en majuscules, dans quatre cas sur sept par rapport au message-

vidéo que le Président de la République prononce lors de la conférence européenne sur l'intelligence artificielle AI Pulse (Fig. 10) :

Figure 10 (Saisie correcte du sigle « IA » par le logiciel d'IA de YouTube) (https://www.elysee.fr/emmanuel-macron/2023/11/17/ai-pulse-la-conference-europeenne-dediee-a-lintelligence-artificielle-le-message-du-president)

Transcription automatique	Transcription révisée
1:13 et c'est dans cette optique que nous 1:16 avons précisément mis en place la 1:19 commission sur l'**IA** et que nous 1:21 accueillerons la prochaine édition du 1:23 sommet sur l'intelligence artificielle à 1:25 Paris l'année prochaine alors je vais […]	1:13 et c'est dans cette optique que nous 1:16 avons précisément mis en place la 1:19 Commission sur l'IA, et que nous 1:21 accueillerons la prochaine édition du 1:23 sommet sur l'intelligence artificielle à 1:25 Paris, l'année prochaine. Alors, je vais […]

Cela est également prouvé par la saisie correcte, même en termes graphiques, de « Open Source » (Fig. 11), figurant dans le même message-vidéo :

Figure 11 (Saisie correcte de la terminologie de l'IA par le logiciel d'IA de YouTube) (https://www.elysee.fr/emmanuel-macron/2023/11/17/ai-pulse-la-conference-europeenne-dediee-a-lintelligence-artificielle-le-message-du-president)

Transcription automatique	Transcription révisée
3:00 continuer d'accompagner je crois aussi 3:02 fortement à l'**Open Source** et nous devons 3:04 en faire une force française nous […]	3:00 continuer d'accompagner. Je crois aussi 3:02 fortement à l'Open Source et nous devons 3:04 en faire une force française. Nous […]

Conclusion

Cette recherche a concerné l'articulation entre les technologies issues de l'intelligence artificielle et l'enseignement/apprentissage des langues étrangères, en vue de mesurer le traitement de l'apport culturel, sur l'exemple de l'enseignement du français langue étrangère, au niveau universitaire de master, filières LANSAD. Parmi les outils de l'intelligence artificielle, nous nous sommes intéressée à la reconnaissance automatique de la parole continue, à l'appui de l'outil de transcription automatique de YouTube, à partir de huit messages-vidéo du Président de la République française prononcés et

publiés entre 2023 et 2024.
L'analyse sur corpus, qui a été menée en classe de FLE pour vérifier la manière dont l'apport culturel est appréhendé par l'outil de transcription automatique de YouTube, témoigne d'une prise en compte problématique et déficitaire de cet apport à plusieurs égards. Lorsque le Président de la République emploie des culturèmes ou choisit de se servir de langues différentes, adaptées au contexte situationnel qui est visé, les résultats du logiciel d'IA de YouTube prouvent que ses performances sont sans aucun doute à améliorer. Ils se traduisent, en effet, par des erreurs par substitution et par omission. Parmi les raisons qui peuvent les expliquer, notre attention a été focalisée sur la variabilité inter- et intra-locuteur, mais également sur l'entraînement de l'algorithme du logiciel en termes de variantes de la langue concernée et de données en entrée, en rappelant que les logiciels de transcription automatique ne sont pas en mesure de prendre en compte la situation d'énonciation. Pour ce faire, nous avons examiné des cas de référents culturels, les « culturèmes » (Pamies, 2017), qui sont cités en langue française par le Président de la République et qui ne sont généralement pas correctement saisis par le logiciel d'IA de YouTube tout en étant bien prononcés, et des formulations composées de mots à charge culturelle partagée (Galisson, 2000) pouvant relever d'une langue autre que le français. Nous avons constaté que même ces derniers sont rarement bien saisis par le logiciel d'IA de YouTube, qui produit des erreurs par substitution ou par omission. Encore, il en va de même pour le choix d'E. Macron de prononcer des mots d'ouverture ou de clôture relevant de la langue du pays qui est concerné par l'événement dont il est question, voire d'entiers passages dans cette langue, comportant un apport culturel fondamental. Or, celui-ci s'avère être annulé lors de la transcription automatique, puisque le contexte n'est pas pris en compte et que d'autres difficultés à bien reproduire ces passages tiennent à des raisons tant techniques, relevant de sa création et de son entraînement, que des variations liées au sujet locuteur. Étant donné que ces passages demandent un travail important de correction en post-édition de la part de nos élèves, ce sont les défauts du logiciel d'IA de YouTube, que nous avons soulignés à côté de ses points de force, qui se transforment en atouts car ils deviennent des observables à étudier, ce qui montre que la gestion numérique de l'aspect culturel est centrale dans l'apprentissage d'une langue étrangère.
Enfin, même le choix des corpus à utiliser s'avère être incontournable afin de bien maîtriser l'apport culturel qui pourrait résulter d'un logiciel de transcription automatique, mais à condition d'avoir d'une vérification humaine des données avant et après que le logiciel performe ses résultats. D'où le rôle, qui reste incontournable, du sujet humain pour bien maîtriser le fonctionnement des logiciels d'IA et pour prendre en compte l'apport culturel, qui fait pour la plupart (encore) défaut lors de la transcription automatique.

Bibliographie

Auzéau, F., Abad, L. (2018). Le corpus : un outil inductif pour l'enseignement-apprentissage de la grammaire. *Synergies France*, 12, 175-187.
Barrault, L. (2008). *Diagnostic pour la combinaison de systèmes de reconnaissance automatique de la parole*. [thèse de doctorat, Université d'Avignon et des Pays du

Vaucluse]. https://theses.hal.science/tel-00424699
Boulton, A. (2007). Esprit de corpus : Promouvoir l'exploitation de corpus en apprentissage des langues. *Texte et corpus*, 3, 37-46.
Chaumartin, F.-R., Lemberger, P. (2020). *Le traitement automatique des langues : comprendre les textes grâce à l'intelligence artificielle*. Malakoff, Dunod.
Develotte, C. (2022). Réflexions sur les changements induits par le numérique dans l'enseignement et l'apprentissage des langues. *Éla. Études de linguistique appliquée*, 160 (4), 445-464.
Fligelstone, S. (1993). Some reflections on the question of teaching, from a corpus linguistics perspective. *ICAME journal*, 17, 97-109.
Galisson, R. (2000). La pragmatique lexiculturelle pour accéder autrement, à une autre culture, par un autre lexique. *Mélanges Crapel*, 25, 47-73.
Haton, J.-P. (2006), *Reconnaissance automatique de la parole. Du signal à son interprétation*. Dunod.
Heba, A. (2021). *Reconnaissance automatique de la parole à large vocabulaire : des approches hybrides aux approches End-to-End*. [thèse de doctorat, Université Paul Sabatier – Toulouse III]. https://theses.hal.science/tel-03616588v1
Lameul, G., Loisy, C. (2014). *La pédagogie universitaire à l'heure du numérique. Questionnement et éclairage de la recherche*. De Boeck Supérieur.
Le Cun, Y. (2019). *Quand la machine apprend : la révolution des neurones artificiels et de l'apprentissage profond*. Éditions Odile Jacob.
Maingueneau, D. (2021). *Discours et analyse du discours : une introduction*. Armand Colin.
Mariani, J. (1990). Reconnaissance automatique de la parole : progrès et tendances. *Journal du Traitement du signal*, 7, 239-266.
Oger, C., Ollivier-Yaniv, C. (2006). Conjurer le désordre discursif : les procédés de « lissage » dans la fabrication du discours institutionnel. *Mots. Les langages du politique*, 81, pp. 63-77.
Pamies, A. (2017). The Concept of Cultureme from a Lexicographical Point of View. *Open Linguistics*, 3, 100–114. https://www.degruyter.com/document/doi/10.1515/opli-2017-0006/html?srsltid=AfmBOoptajLLXLkSjvCv0LS-4syeCUh6rn-6FfaVPsNWUXnpQtnus4dp
Puren, C. (2022). Innovation didactique et innovation technologique en didactique des langues-cultures : approche historique. *Recherche et pratiques pédagogiques en langues*, 41/1. https://doi.org/10.4000/apliut.9708
Rastier, F. (2021). Data vs corpora. Dans : D. Mayaffre et L. Vanni (dir.), *L'intelligence artificielle des textes : des algorithmes à l'interprétation* (p. 203-246). Honoré Champion.
Raus, R., Silletti A.M. (2022). Expérimentations pédagogiques. *De Europa. European and Global Studies Journal*, Special Issue 2022, Multilinguisme et variétés linguistiques en Europe à l'aune de l'intelligence artificielle, 199-214. https://www.collane.unito.it/oa/items/show/132
Raus, R., Tonti, M. (2025). Intelligence artificielle, corpus et diversité linguistique : enjeux et perspectives. Introduction. *Langages*, 237 (1), 7-20.
Silletti, A. M. (2023). La transcription automatique comme outil d'apprentissage du FLE :

quelques réflexions à partir d'un corpus oral. *Studi di Glottodidattica*, 2, 41-52.
Silletti, A. M. (2025). La transcription automatique au service de l'apprentissage des langues : quelques réflexions sur la phrase complexe en français. *Langages*, 237 (1), 131-152.
Silletti, A. M. (soumis). *Complexité morphosyntaxique de la langue française et apport de la transcription automatique en didactique du FLE/FOS*.
Tancoigne, É., Corbellini, J.-P., Deletraz, G., Gayraud, L., Ollinger, S. et Valero, D. (2020), *La transcription automatique : un rêve enfin accessible ? Analyse et comparaison d'outils pour les SHS. Nouvelle méthodologie et résultats*. [rapport de recherche – MATE-SHS]. halshs-02917916v2

Corpus

Message du Président de la République pour le Forum mondial de l'eau de Bali (5min39) – 21 mai 2024, https://www.youtube.com/watch?v=8Lzyg43gBm8
50e anniversaire de la Révolution des Œillets : le message du Président Emmanuel Macron (3min42) – 24 avril 2024, https://www.youtube.com/watch?v=evKMDUa9EjE
Relève franco-britannique de la garde à l'occasion des 120 ans de l'Entente Cordiale (3min30) – 8 avril 2024, https://www.youtube.com/watch?v=AAqXumWp6ic
80e anniversaire de la Libération : le message du Président Emmanuel Macron (5min30) – 6 mars 2024, https://www.youtube.com/watch?v=fMrss0S0YKY
Sixième édition de la Conférence Nationale Humanitaire (6min24) – 19 décembre 2023, https://www.youtube.com/watch?v=aNJK22D95Ro
Journée internationale de lutte contre les violences faites aux femmes : le message du Président Emmanuel Macron (6min56) – 25 novembre 2023, https://www.youtube.com/watch?v=GsyQzTtm1PU
ai-PULSE, la conférence européenne dédiée à l'intelligence artificielle : le message du Président (8min19) – 17 novembre 2023, https://www.youtube.com/watch?v=QEyG-Xkm9EE
150 ans de l'Institut de droit international : le message du Président Emmanuel Macron (5min35) – 28 août 2023, https://www.youtube.com/watch?v=K3DBYafYpHo

---

1 Dans les exemples, le jaune est utilisé pour indiquer les modifications qui sont apportées en post-édition, tandis que le gras et le soulignement signalent le cas dont il est question dans la transcription automatique.

2 Pour écouter les passages qui font l'objet de nos exemples, il suffit de cliquer sur le lien et de se servir des balises temporelles accompagnant chaque transcription pour repérer chaque passage.