

# Robots conversationnels en classe de philosophie : une prise de conscience possible

Etude de cas.

**Inès Drège**

professeur de philosophie

Lycée Henri Parriat (Montceau-les-Mines)



**Descriptif :**

*Cette contribution est issue d'une présentation faite le 10 juin 2024 au séminaire SIAFO du Gis2if.*

**Résumé**

*J'enseigne la philosophie en terminale depuis une dizaine d'années ; je vais présenter les questionnements qui m'ont conduit à introduire les IA génératives en classe ; puis je présenterai quelques situations concrètes en classe ; et je terminerai par les questionnements que cette nouvelle situation suscite.*

## Questionnement initial

La popularité des IA génératives a réactivé un questionnement que j'avais déjà concernant les langues : la diversité des langues de l'école, des langues des élèves, et maintenant de la langue de la machine qui prétend prendre en charge la langue de tous, dans un système scolaire français très monolingue. Les enseignants de philosophie voudraient faire reconnaître aux élèves qu'un écrit rédigé en première personne est la condition de l'élaboration d'une pensée instruite et raisonnée ; mais si des machines grand public écrivent désormais mieux que la plupart d'entre nous, les élèves sont en droit de nous demander pourquoi nous leur demandons de passer quatre heures à rédiger des dissertations avec un simple papier et un stylo. Ils savent d'ailleurs spontanément que la production écrite est l'exercice qui permet au système scolaire de distinguer entre ceux qui ont le capital linguistique pour réussir leurs études, et ceux qui ne l'ont pas ; Bourdieu a montré l'articulation de ce capital à la trajectoire sociale des étudiants dans le processus de sélection scolaire<sup>1</sup>.

---

<sup>1</sup> Bourdieu, P. & Passeron, J.C. (1970). *La reproduction*. Les éditions de Minuit. p. 91-92

La première question qui m’a fait entrer sur ces sujets était : pourquoi la langue des agents conversationnels exerce-t-elle un tel attrait sur les élèves ? Le texte produit par l’IA confine à une possibilité menaçante dont parle Derrida, en réponse au déracinement linguistique qui a été le sien – Juif d’Algérie, il n’a jamais appris l’arabe ou l’hébreu et a appris le français dans sa version scolaire en Algérie, une langue qu’il dit n’être pas sa langue maternelle bien qu’il n’en ait pas eue d’autre, langue du colon dont il a hérité en étant séparé par une mer des modèles linguistiques et culturels qui nourrissaient le français métropolitain. Celui qui ne dispose pas d’une langue d’identification stable, dit Derrida dans *Le monolinguisme de l’Autre*, risque d’être conduit à la recherche « de stéréotypes homogènes et conformes au modèle français « moyen » ou dominant, une (...) amnésie sous la forme intégrative : une (...) folie »<sup>2</sup>. C’est à peu de chose près la réponse de nivellement linguistique mondialisé que semblent proposer les grands modèles de langage ; c’est une menace à laquelle les élèves sont particulièrement exposés, eux qui sont en train d’acquérir la langue des matières et sont dans une situation de diglossie par rapport à ces langues.

L’objectif pédagogique, quand j’ai décidé d’introduire les outils dans mes classes, c’était donc de montrer à mes élèves que l’IA n’allait pas écrire à leur place et qu’ils avaient besoin de conquérir leur parole, fût-elle en décalage par rapport à la norme linguistique qu’on leur demande d’atteindre à l’école. Je voulais leur montrer qu’ils n’ont pas à rougir de leur façon d’écrire : les meilleures dissertations qu’on corrige au baccalauréat ne sont pas écrites dans une langue standardisée, portent la marque de la subjectivité de l’auteur, et même commettent des erreurs.

### Cas pratiques en classe

Il a déjà fallu montrer à mes élèves que quand nous lisons une copie plagiée sur une IA, en tant qu’enseignants, nous identifions les « stéréotypes homogènes et conformes au modèle français « moyen » ou dominant » dont parle Derrida. J’ai photocopié une copie dans laquelle l’élève s’était contentée de recopier ce que proposait Chat-GPT sur le sujet « une société sans religion est-elle possible ? », l’ai distribuée en classe et ai demandé aux élèves de se mettre en situation de correction, sans dire l’origine de l’écrit qu’ils corrigeaient. Ils croyaient corriger un camarade ; ils ont été indulgents. Quand je leur ai dit que la copie avait été intégralement produite par une IA générative, ils m’ont répondu que c’était impossible parce que la copie portait des erreurs syntaxiques ; l’élève les avait introduites en recopiant. J’ai mesuré que mes élèves ne voyaient pas du tout qu’ils avaient affaire à un écrit standardisé.

Dans l’enseignement de la dissertation, il faudrait questionner l’impact de la « méthodologie » sur la capacité des élèves à s’approprier l’exercice : ils appliquent souvent ses différentes étapes

---

<sup>2</sup> Derrida, J. (1996). *Le monolinguisme de l’Autre*, Galilée, p. 116

comme on réaliserait une recette de cuisine, et ne comprennent pas ensuite pourquoi ce qu'ils nous ont rendu n'est pas à proprement parler une dissertation mais plutôt un simulacre de raisonnement formellement équilibré. Ce que produisent les agents conversationnels leur semble alors tout à fait conforme aux attendus de l'enseignant. Quand j'ai proposé aux élèves d'utiliser les outils en étayage de la rédaction d'une dissertation en binômes en mars 2023, nous avons porté ce regard critique sur la langue de la machine. Je les accompagnais dans la lecture et leur montrais la tendance qu'elle avait à produire des rythmes ternaires par exemple, l'uniformité du nombre de caractères dans les phrases.

En même temps, je leur disais qu'on attendait autre chose d'eux, et bien plus finalement. J'insiste : j'ai beaucoup d'élèves qui en sont à me demander combien il faut écrire de lignes par paragraphe, comme si la qualité d'une argumentation résidait dans le simple équilibre formel du texte. La plupart des groupes m'a rendu un travail manuscrit ; ils ont été auteurs de l'argumentation et se sont affranchis de la langue de la machine. En les corrigeant, les passages des écrits que j'ai soulignés comme étant les moins efficaces étaient ceux où ils s'étaient contentés de recopier le texte de l'IA générative.

Il faudrait vérifier la plus-value de l'outil pour faire progresser des élèves sur le temps long, pour chaque compétence mobilisée dans la production écrite d'une dissertation, en tenant compte du fait qu'écrire est un *processus* très complexe et à mille lieues de la production instantanée de l'IA générative : pour la simple compétence désignée dans le programme comme « circonscrire les questions dont la réponse demande une réflexion préalable » par exemple, il faudrait voir si l'interaction avec l'agent conversationnel leur permet de progresser davantage que nos prescriptions méthodologiques<sup>3</sup>.

En face de machines qui produisent instantanément un *résultat* écrit complexe, on doit tenir compte du fait qu'écrire est une *tâche*, un *processus* complexe au cours duquel l'élève se questionne, et qui implique toute l'épaisseur d'une scolarité pour être exécuté. Autrement dit, il faudrait partir de la complexité des processus cognitifs impliqués dans la rédaction pour vérifier si l'étayage de l'agent conversationnel les fait progresser ou non. En tant qu'enseignants, on ne peut pas seuls vérifier la plus-value de ces outils pour les progrès de nos élèves.

Une étude américaine effectuée par des chercheurs en informatique montre que les grands modèles de langage échouent aux tests de l'écriture créative<sup>4</sup> auxquels les écrivains réussissent. Les élèves sont généralement peu créatifs dans leurs écrits scolaires : ils y voient à juste titre un

---

<sup>3</sup> Et il faudrait étendre une telle recherche à toutes les compétences mobilisées dans la production d'écrits aussi complexes que la dissertation de philosophie : originalité des exemples convoqués en illustration de la pensée conceptuelle ; organisation et confrontation des points de vue sur un problème déterminé ; rigueur de la démonstration ; pertinence des choix de l'élève dans la mobilisation des connaissances issues de sa culture philosophique ; élaboration des concepts ; réflexion sur la polysémie des notions pour articuler une réponse...

risque d'être à côté de la norme attendue par l'école. Pourtant la dissertation est un écrit créatif ; si nous échouons à transformer la représentation que les élèves ont de l'écrit scolaire comme un écrit formel et standardisé, on s'expose à une explosion des inégalités sociales, car pour être créatif il faut s'être confronté à l'altérité et avoir de la culture. J'enseigne dans le lycée qui obtient les plus faibles résultats au baccalauréat de l'académie de Dijon, et je sais que nos exercices, actuellement, renforcent les inégalités sociales.

Malgré nos efforts pour valoriser dans l'évaluation tout ce qui ressemble à un effort de l'intelligence, les élèves retiennent de la dissertation que si c'est en trois parties et trois sous-parties que l'enseignant évalue leur effort d'invention, c'est ainsi qu'ils devraient penser et écrire. Les élèves sont tellement en difficulté dans cet exercice qu'au baccalauréat, on encourage au maximum la restitution du cours : si c'est ça faire une dissertation, elle sera mieux exécutée par la machine.

Cette année, face au risque de triche, j'ai systématiquement évalué le brouillon, puisque dans l'élaboration d'une dissertation c'est le geste le plus créatif et qui ne peut pas être produit par une IA. Bien trop peu d'élèves élaborent leur pensée au brouillon ; on ne parvient pas à l'enseigner comme un prérequis. Je présente la dissertation comme une fiction dont l'ordre de la présentation est complètement disjoint de l'ordre d'invention : au brouillon, l'effort d'invention se fait dans le désordre ; sur la copie, l'élève choisit un ordre fictif pour proposer un itinéraire de la prétendue découverte d'un problème à sa résolution en trois parties antagonistes. Personne ne pense ainsi ; personne n'écrit ainsi, sauf à l'école.

Je voulais valoriser la langue des élèves face à la langue de la machine : j'ai paramétré un agent conversationnel qui avait comme source un lexique d'argot adolescent recueilli par des lycéens de la région parisienne pendant la pandémie de covid. Pendant un cours, j'avais demandé à une classe de choisir un sujet de débat en lien avec le cours sur la justice : j'ai proposé aux élèves de l'utiliser pour trouver des idées. C'est Rahid, qui se revendique comme un utilisateur expert des IA, qui s'est proposé d'interagir avec la machine ; il était fier de montrer qu'il savait nous donner matière à réflexion en s'appuyant sur la machine. Au bout de quelques interactions, Rahid s'est exclamé : « ça ne marche pas Madame votre truc, il parle comme en 2017 ! » Il serait intéressant de faire produire aux élèves un recueil d'un lexique adolescent en constante évolution, pour créer un agent conversationnel au plus près de leur langue qui leur montrerait que le texte de la machine est figé et incapable d'inventer la prochaine expression à la mode, alors que les jeunes sont sans cesse en train d'inventer leur langue.

---

<sup>4</sup> ["Art or Artifice? Large Language Models and the False Promise of Creativity"](#), Tuhin Chakrabarty, Philippe Laban, Divyansh Agarwal, Smaranda Muresan, Chien-Sheng Wu, 2024

## Nouveaux questionnements

L'usage des IA génératives conduira probablement à une perte du prestige social de l'écriture, ou au moins d'un certain type d'écriture à laquelle nous formons la jeunesse à l'école. Si l'exigence de rigueur continue à être valorisée par les enseignants au détriment de l'effort d'invention des élèves, si par ailleurs les normes académiques de l'excellence en écriture se transforment sous l'effet de la massification de l'usage des grands modèles de langage, les générations d'élèves à venir sont exposées à un double risque. Celui d'un nivellement par le bas de nos exigences, et d'une pensée globalement appauvrie par des exercices qui n'encouragent pas suffisamment la tentative mais sanctionnent l'erreur n'est pas à exclure.

On peut également craindre une aggravation profonde des inégalités scolaires si l'école n'est pas capable de changer ses exercices et ses pratiques d'évaluation, inégalités qui seront favorables à ceux qui auront produit des écrits qui encouragent l'effort d'invention et la capacité de nuancer son propos, en s'appuyant sur des compétences de maîtrise de la langue dans sa complexité sémantique et syntaxique qui ne peuvent être acquises qu'indépendamment de l'usage des IA génératives.

Éviter ce problème suppose, à mon avis, un travail d'écriture qui inclurait les IA génératives de façon précoce, dès l'école élémentaire : il s'agirait de proposer aux élèves des exercices qui, tout en travaillant les compétences fondamentales de la maîtrise de la langue, développeraient les capacités de métacognition qui sont à l'œuvre dans tout processus de rédaction. Actuellement, pour le premier degré, à ma connaissance les outils disponibles se contentent de proposer des exercices renforçant des compétences déjà acquises dans le contexte de l'enseignement ordinaire.

Or il me semble que le potentiel de l'IA pour l'enseignement est dans la capacité de la machine à rendre la langue de l'école accessible (séquençage des consignes, traduction, accès au sens facilité par l'interaction homme-machine, ...). Les grands modèles de langage les plus populaires permettent donc déjà d'exploiter ce potentiel, mais le dispositif pédagogique à mettre en place peut être périlleux : la crainte d'un manque d'investissement des élèves dans la tâche, quand nous introduisons les outils numériques en classe, peut s'avérer légitime quand les élèves sont très jeunes ; pour des enseignants, laisser des élèves du premier degré, ou même de collège, jouer avec du texte sur un grand modèle de langage, est difficile. La maturité des lycéens est un facteur facilitant.

La philosophie est une discipline dans laquelle il est difficile de faire l'économie des devoirs maison, mais il va nous falloir imiter nos collègues de mathématiques au moment de l'apparition des calculatrices et trouver davantage de stratégies d'évaluation du *processus* d'élaboration de la pensée argumentative, quitte à minorer franchement l'évaluation du résultat.

Pour clore sur une perspective plus optimiste, la popularité des grands modèles de langage peut aussi favoriser les apprentissages sauvages, y compris chez les publics les plus fragiles socialement. C'est une question qui me tient à cœur parce qu'un autre versant de mon activité professionnelle concerne la scolarisation des élèves issus de familles itinérantes et de Voyageurs ; la proportion d'illettrisme est très élevée parmi les gens du voyage. Aujourd'hui, grâce aux outils numériques ces publics peuvent faire lire leurs documents administratifs sans l'intervention humiliante d'une assistante sociale ; il y a quelques années, une telle situation était inimaginable. Peut-on s'appuyer sur les conclusions actuelles des chercheurs en science du langage pour anticiper une démocratisation de la production écrite pour des publics éloignés de l'école ? Il faudrait observer les nouvelles manières d'entrer dans l'écriture en comparant la situation actuelle avec celle des migrants apprenant la langue du pays d'exil avec les outils de traduction automatique.

La recherche montre que les outils de traduction automatique ne sont efficaces pour écrire dans une langue seconde qu'à un certain niveau d'autonomie linguistique<sup>5</sup> ; quelles sont les compétences du scripteur qui pourraient être renforcées par l'étayage d'un grand modèle de langage, et dans quelle situation pédagogique ? Comment l'utilisation d'un grand modèle de langage pourrait-elle appuyer la production écrite chez les personnes ayant une compétence limitée en écriture ?

Si on est optimiste, la perspective pourrait être la suivante : la popularité des IA génératives pourrait provoquer simultanément un déclin de nos exercices scolaires, mais encore faudra-t-il que l'école sache s'adapter, et une démocratisation de la production d'écrits indispensables pour des publics illettrés.

---

<sup>5</sup> Bozhinova, K., Narcy-Combes, J., Mabrou, A. (2020)., "[Ecrire en langue additionnelle : un besoin de complexifier les modèles](#)"