Banque BCPST Inter-ENS/ENPC/Mines - Session 2023

RAPPORT
ÉPREUVE ÉCRITE D'ANGLAIS

Écoles concernées : ENS (Paris) – ENS de Lyon – ENS Paris-Saclay - ENPC - Mines

JURY : Gabriel Lattanzio

Coefficient :

(en pourcentage du total d'admission)

ENPC/Mines : 3,8%
ENS Lyon : 2,5%
ENS Paris-Saclay : 3,1%
ENS Paris-Ulm : 2,1%

1. Présentation de l'épreuve et du sujet

Format de l'épreuve

12 points de version, à partir d'un document de presse écrite, sur l'actualité scientifique en lien avec une question de société. 8 points de rédaction, sur des questions dérivées du texte.

Document étudié

Le jury a fait travailler les candidats et les candidates sur un article de la revue *Nature*, numéro 614 en 2023, pages 214 à 216 : *What ChatGPT and generative AI mean for science*. Ce texte a été choisi en hiver, donné aux candidats au printemps et corrigé au début de l'été. Il est remarquable qu'en l'espace de quelques mois, la conversation collective sur l'intelligence artificielle a progressé presque aussi rapidement que la technologie concernée. L'auteur de cet article mettait en garde sur les limites de formes d'intelligence artificielle à des fins de recherche scientifique. Comme l'année dernière, nous avons voulu que ce sujet permette aux candidats de produire les réflexions les plus riches possibles. Le document était une invitation à réfléchir sur les conditions de production de la science et sur le support même de la pensée, le langage.

Sur l'exercice et ses objectifs

Le travail de traduction est une façon efficace de mesurer les compétences langagières des candidats. Il permet de mesurer la capacité à la compréhension du texte, mais aussi une appréciation de la façon dont la pensée est formulée de façon naturelle dans la langue-cible, en l'occurrence le français. A ce stade de leur formation, les candidats doivent démontrer que leur pratique de l'anglais a dépassé l'apprentissage du vocabulaire usuel courant et des champs lexicaux de leur spécialité ; ils doivent également faire la preuve de leur capacité à s'exprimer d'une façon idiomatique. Le sens n'est plus déduit par le lecteur. Il est exact. Quant à la partie argumentative, le candidat met en évidence sa capacité à convaincre. Il est toujours appréciable quand un candidat prouve sa capacité à prendre en compte des points de vue contraires, mais ce n'était pas une attente particulière.

2. Un profil-type du candidat 2023, et ce que les candidats de 2024 devront améliorer.

Nous devons confier avoir souri à la lecture de plusieurs copies. Les candidats ont dans l'ensemble su décrire les failles de l'intelligence artificielle. Cette dernière se répète, elle mobilise des idées prises aléatoirement au pot commun d'internet, et elle privilégie le "stylistiquement plausible" à ce qui serait un propos dont la logique est irréprochable. L'intelligence naturelle ou culturelle de certains candidats présentait quant à elle les défauts suivants : des répétitions, des idées aléatoirement prises au pot commun d'internet, et une préférence allant au stylistiquement

plausible davantage qu'à une logique irréprochable. Nous voyons peut-être plus facilement dans la machine des failles qui nous sont tout autant propres.

Le profil-type du candidat de 2023 est quelqu'un qui a lu assez pour savoir produire un écrit assez sophistiqué pour être classé B2 selon les normes du Cadre européen commun de référence pour les langues, tout en étant capable de fautes très simples, habituellement surmontées dans le passage du niveau A2 au niveau B1. Nous comprenons cela comme la conséquence d'un système d'évaluation commun dans l'enseignement secondaire : un bon élève en anglais a tendance à privilégier la longueur de sa rédaction à l'exactitude de la grammaire. Parce que les candidats sont aujourd'hui déjà capables d'écrire beaucoup et avec une relative sophistication, il est important d'orienter ses efforts vers l'exactitude de la langue, qui est une ambition bien moins satisfaite aujourd'hui. L'intelligibilité et la compréhension sont des compétences bien acquises, il faut désormais se concentrer sur la qualité.

Nous notons cette année par rapport à l'année dernière, une plus grande maîtrise d'expressions authentiques, mais un registre de langue qui n'est pas toujours adapté. Cela est le produit d'un apprentissage de l'anglais résultant d'une forte exposition à la langue par les productions culturelles communes, mais cela ne convient pas pour une ambition académique. Les candidats doivent distinguer l'anglais vernaculaire et le geste professionnel. Cela ne veut pas dire faire l'économie d'effets, mais il faut trouver un équilibre entre l'austère et le familier. La lecture même de *Nature* montre bien que l'anglais permet cet équilibre.

Insistons également sur l'importance de présenter une copie lisible. Ce n'est pas la responsabilité du jury de deviner ce qui est écrit. Le bénéfice ne peut aller à l'élève si le mot est incompréhensible. Certaines copies sont raturées toutes les lignes. Une copie en particulier a même trois ratures par ligne. Il faut traiter sa copie comme un bel objet que l'on rend.

3. Correction de la langue anglaise

L'emploi d'un anglais riche est valorisé et on peut saluer les efforts de nombreux candidats et de nombreuses candidates à cet égard. À des fins de correction, nous compilons ici les erreurs les plus fréquentes et remarquons qu'elles correspondent à des erreurs habituelles dans l'apprentissage de l'anglais en France.

Idiomatisme

Le texte à traduire cette année n'était pas particulièrement technique. Il ne nécessitait qu'une connaissance relativement minimale de termes scientifiques, professionnels et technologiques. Sa difficulté résidait davantage dans des qualités propres à la langue anglaise : sa plasticité, son inventivité, sa qualité synthétique. La lecture de *Nature* peut surprendre tant leur rigueur journalistique ne signifie pas qu'ils s'en tiennent à une rédaction austère. La façon par laquelle l'anglais peut associer des mots, la façon dont la langue se prête aux néologismes immédiatement identifiables par les locuteurs natifs est aussi une difficulté que tous les candidats n'ont pas su surmonter. L'exemple du texte qui en témoigne était *much-hyped generative AI chatbot-style tools*. Nous invitons donc les candidats à se préparer à un travail de déduction sur les néologismes, et sur la portée des adjectifs et des adverbes.

Syntaxe et grammaire

En ce qui concerne le travail de rédaction personnelle, nous avons retrouvé les fautes habituelles de celles et ceux dont le niveau oscille entre les normes A2 et B1 du CECRL. Les candidats et les candidates au niveau général B2 pouvaient également faire certaines des fautes suivantes : la conjugaison du présent simple, l'emploi excessif du déterminant *the*, l'oubli ou l'ajout inutile de *to*, des formes négatives maladroites (causé par le recours excessif aux

contractions comme sur *don't*), la confusion entre dénombrables et indénombrables, la modalisation et son intégration dans la conjugaison, une ignorance des verbes irréguliers, entre autres erreurs.

Vocabulaire

On attend des candidats et des candidates qu'ils et elles aient acquis un vocabulaire suffisamment riche pour pouvoir s'exprimer de manière nuancée et claire. Les gallicismes sont encore trop fréquents et sont l'indicateur le plus certain d'une faible exposition à la langue anglaise (*scientifics* au lieu de *scientists*). Au-delà du lexique, il est important de maîtriser les expressions consacrées en anglais et de ne pas en inventer. Les candidats et les candidates en difficulté ont parfois l'impression qu'ils peuvent compenser leur ignorance du lexique par une structure grammaticale complexe. Il faut en réalité faire exactement l'inverse. Plutôt qu'un vocabulaire simple associé à une syntaxe complexe, il faut privilégier un vocabulaire riche associé à une syntaxe simple.

4. Correction de la langue française

Trop souvent, le niveau de français ne satisfait pas les attentes légitimes à ce stade de la formation des candidats. Citons ces exemples : "été" au lieu de "était", "mettent en exerguent", "une intelligence basé", "chaque expériences", "peut importe", "chaques possibilités". Toutes les fautes possibles ont été écrites. L'accord des adjectifs, la conjugaison, la différence entre "a" et "à", des pluriels aléatoires, des accents avant des consonnes doublées, et tant d'autres.

Pour un autre niveau d'exigence, notons, entre autres exemples, la phrase suivante qu'il fallait traduire : "researchers emphasize that LLMs are fundamentally unreliable at answering questions". Trop de candidats ont essayé de produire une traduction en partant du verbe "are", et se sont sentis obligés de coller à l'ordre des mots tels qu'ils sont en anglais. Ils ne savaient plus quoi faire du début de la phrase avec "emphasize". Il faut comprendre, puis essayer de produire une phrase en français qui serait naturelle. Il est particulièrement important que les candidats apprennent l'ordre dans lequel apparaissent les adjectifs autour d'un substantif. Contrairement à l'anglais, ils sont plus souvent après le nom, contrairement à l'anglais, pour lequel souvent, ce qui précise précède.

5. Appréciation générale

Bien que produites à partir de deux exercices distincts, les notes correspondent assez fidèlement aux exigences définies par le Cadre européen commun de référence pour les langues, avec pour équivalence générale une note autour de 7 pour un niveau A2 moyen, autour de 12 pour un B1 solide, autour de 16 pour un B2 avec un nombre raisonnable de fautes. S'y ajoutent des variations selon la capacité du candidat à avoir satisfait les exigences de l'exercice en lui-même, et selon des contre-sens plus ou moins graves dans la traduction. Le jury s'attendait à un niveau plus homogène, moins prononcé aux extrêmes. Nous saluons le travail des meilleurs candidats et candidates. À leur érudition s'ajoutait de belles tournures idiomatiques, dont l'utilisation n'était pas accessoire mais permettait de modaliser à bon escient leur propos. Il est remarquable que les performances dans les deux exercices, la traduction et l'argumentation, illustrent une corrélation entre les deux compétences. Les meilleurs traducteurs sont souvent les plus convaincants, les plus à même de proposer une analyse personnelle. La nuance dans la compréhension du propos d'autrui est un indicateur de la capacité à nuancer sa propre expression. Nous sommes heureux de pouvoir témoigner qu'il existe une cohorte de jeunes scientifiques qui ont démontré leur capacité à écrire en anglais. Ils feront de cette compétence un grand atout de leur évolution professionnelle.

6. Sujet 2023

ANGLAIS

I. VERSION (12 points, titre à traduire également)

What ChatGPT and generative AI mean for science

Researchers are excited but apprehensive about the latest advances in artificial intelligence.

In December, computational biologists Casey Greene and Milton Pividori embarked on an unusual experiment: they asked an assistant who was not a scientist to help them improve three of their research papers. Their assiduous aide suggested revisions to sections of documents in seconds; each manuscript took about five minutes to review. In one biology manuscript, their helper even spotted a mistake in a reference to an equation. The trial didn't always run smoothly, but the final manuscripts were easier to read — and the fees were modest, at less than US$0.50 per document.

This assistant is not a person but an artificial-intelligence algorithm called GPT-3, first released in 2020. It is one of the much-hyped generative AI chatbot-style tools that can churn out convincingly fluent text, whether asked to produce prose, poetry, computer code or to edit research papers. The most famous of these tools, also known as large language models, or LLMs, is ChatGPT, a version of GPT-3 that shot to fame after its release in November last year because it was made free and easily accessible. "I'm really impressed," says Pividori, who works at the University of Pennsylvania in Philadelphia. "This will help us be more productive as researchers."

But researchers emphasize that LLMs are fundamentally unreliable at answering questions, sometimes generating false responses. ChatGPT and its competitors work by learning the statistical patterns of language in enormous databases of online text — including any untruths, biases or outmoded knowledge. When LLMs are then given prompts, they simply spit out, word by word, any way to continue the conversation that seems stylistically plausible.

The result is that LLMs easily produce errors and misleading information, particularly for technical topics that they might have had little data to train on. LLMs also can't show the origins of their information; if asked to write an academic paper, they make up fictitious citations. "The tool cannot be trusted to get facts right or produce reliable references," noted a January editorial on ChatGPT in the journal Nature Machine Intelligence. With these caveats, ChatGPT and other LLMs can be effective assistants for researchers who have enough expertise to directly spot problems or to easily verify answers. But the tools might mislead naive users. (...)

Adapted from *Nature*, Vol 614, 9th of February 2023, Pages 214-216.
https://www.nature.com/articles/d41586-023-00340-6

II. QUESTIONS (8 points, minimum de 100 mots par question)

1. What are reasons why Large Language Models fail to be perfectly accurate?

2. Do you believe that artificial intelligence will revolutionize scientific work? How do you envision responsibilities to be shared between human and machine?