
L'exposé oral en langue étrangère à l'ère de l'intelligence artificielle : vers une refonte des critères d'évaluation ?

Mathis Donzé
mathis.donze@hep-bejune.ch

Directrice : Mme Sheila Padiglia

Travail Écrit de Recherche
Master of Advanced Studies en enseignement
Disciplines d'études : anglais et histoire (secondaire I et II)

REMERCIEMENTS

Je souhaite exprimer ma plus sincère gratitude à ma directrice de Travail Écrit de Recherche, Madame Padiglia, pour ses conseils avisés, sa disponibilité et son accompagnement tout au long de l'élaboration de cette étude.

J'adresse également mes chaleureux remerciements à ma mentor, Madame Trapletti, pour la qualité de nos échanges lors de nos rendez-vous hebdomadaires, qui ont contribué à enrichir et à affiner ma pratique enseignante.

Bien que les chances qu'ils lisent un jour ces lignes restent statistiquement assez faibles, une pensée reconnaissante va également à mes élèves, pour leur implication dans ce travail, leur sérieux et leur bonne volonté tout au long du dispositif.

Enfin, je remercie vivement mon frère, Nils, pour sa relecture attentive, qu'il m'a généreusement accordée malgré une période particulièrement intense à l'EPFL.

RÉSUMÉ

À l'ère de l'intelligence artificielle générative (IAG), les modalités de production des tâches scolaires, notamment les exposés oraux en langue étrangère, connaissent une transformation rapide qui interroge la validité de certaines pratiques évaluatives. Cette recherche examine la pertinence des critères d'évaluation traditionnels de l'exposé oral lorsque les élèves peuvent recourir à des outils d'IAG pour préparer leur discours. L'étude repose sur une recherche empirique menée dans deux classes d'anglais du secondaire I en Suisse romande. L'analyse est fondée sur un dispositif comparatif : un exposé préparé sans IA, puis un exposé, de même ampleur, avec usage autorisé mais encadré de l'IA. Les données croisent les notes des performances, les auto-évaluations des élèves et les usages déclarés de l'IA. Les résultats mettent en évidence des écarts de notes globalement modestes entre les deux exposés, ainsi qu'une dissociation entre performance observée, sentiment de progression orale et appropriation déclarée du discours. Les conclusions suggèrent également l'intérêt, dans ce contexte précis, d'une grille d'évaluation « IA-informée » valorisant davantage l'appropriation, les réponses personnelles et le processus. Plus largement, l'étude souligne la nécessité de repenser la validité, l'équité et même le sens de l'évaluation de l'oral à l'ère des outils génératifs.

Mots-clés : intelligence artificielle générative (IAG), évaluation, exposé oral, langue étrangère, secondaire I.

TABLE DES MATIÈRES

INTRODUCTION.....	6
1. PROBLÉMATIQUE ET CADRE THÉORIQUE	8
1.1. POINT DE DÉPART DE LA RÉFLEXION : L'IA ET LA FRAGILISATION DE L'INFÉRENCE ÉVALUATIVE.....	8
1.2. L'EXPOSÉ ORAL EN LANGUE ÉTRANGÈRE : UNE TÂCHE COMPLEXE AU SERVICE DE COMPÉTENCES MULTIPLES..	8
1.3. IAG ET « EFFET DE SURFACE » : VERS UN PAROXYSMES DE LA DÉCORRÉLATION FORME-COMPRÉHENSION ?	9
1.4. VALIDITÉ ET ÉQUITÉ : DEUX EXIGENCES MISES SOUS TENSION	9
1.5. VERS UNE ADAPTATION RAISONNÉE DES CRITÈRES : JUSTIFICATION D'UNE GRILLE « IA-INFORMÉE »	10
1.6. QUESTION DE RECHERCHE ET OBJECTIFS DE L'ÉTUDE	11
2. MÉTHODOLOGIE.....	12
2.1. DISPOSITIF DE RECHERCHE : LES GRANDES LIGNES	12
2.2. LA NATURE DES DONNÉES ET LEURS PROCÉDURES RESPECTIVES DE RECUEIL.....	12
2.3. LES MÉTHODES D'ANALYSE DES DONNÉES.....	14
3. PRÉSENTATION ET ANALYSE DES RÉSULTATS	15
3.1. NATURE DES MESURES ET PRÉCAUTIONS ANALYTIQUES	15
3.2. ÉVOLUTION DES PERFORMANCES : COMPARAISON DES NOTES (SANS IAG VS AVEC IAG).....	15
3.3. EFFET DE LA GRILLE : COMPARAISON NOTE IA-INFORMÉE VS RE-COTATION AVEC LA GRILLE STANDARD	16
3.4. ANALYSE DES AUTO-ÉVALUATIONS INDIVIDUELLES (Q1–Q11) : ÉVOLUTIONS GÉNÉRALES	17
3.5. ANALYSE PAR OPTION SUR LES ITEMS LES PLUS LIÉS À LA PROBLÉMATIQUE (Q3, Q7, Q8, Q11)	18
3.6. ANALYSE DES ITEMS SPÉCIFIQUES À L'IA (Q12–Q21) : USAGE, VÉCU, DÉLÉGATION	19
3.7. NOTES VS PERCEPTIONS : CONVERGENCES ET ÉCARTS ATTENDUS	20
4. INTERPRÉTATION ET DISCUSSION	21
4.1. STABILITÉ DES PERFORMANCES ET ROBUSTESSE RELATIVE DE LA TÂCHE D'EXPOSÉ ORAL.....	21
4.2. ÉCART ENTRE PERFORMANCE ÉVALUÉE ET APPROPRIATION DÉCLARÉE : UN RÉSULTAT CENTRAL.....	21
4.3. L'IAG COMME OUTIL D'EFFICACITÉ ET D'ALLÈGEMENT DE LA CHARGE COGNITIVE	22
4.4. AUTO-ÉVALUATION, SENTIMENT DE MAÎTRISE ET POTENTIELLE ILLUSION D'EFFICACITÉ	23
4.5. L'EFFET DE LA GRILLE IA-INFORMÉE ET LA QUESTION DE LA VALIDITÉ ÉVALUATIVE	23
4.6. TRANSFORMATION DU RAPPORT À L'EFFORT, AU TEMPS ET AU SENS DE L'APPRENTISSAGE.....	24
4.7. ENJEUX ÉTHIQUES ET PROFESSIONNELS POUR L'ENSEIGNEMENT AU SECONDAIRE I.....	24
4.8. OUVERTURE RÉFLEXIVE : AGENTIVITÉ, JUGEMENT ET FINALITÉS ÉDUCATIVES À L'ÈRE DE L'IA.....	24
CONCLUSION	26
BIBLIOGRAPHIE	29
ANNEXES	31

LISTE DES ABRÉVIATIONS

- HEP-BEJUNE : Haute École Pédagogique BEJUNE ;
- TER : travail écrit de recherche ;
- PER : plan d'études romand ;
- O2 / O3 / O4 : option 2 / 3 / 4 (système, sous peu caduc, de répartition des élèves au secondaire I dans le canton du Jura) ;
- IA(G) : intelligence artificielle (générationnelle) ;
- Q1–Q21 : items du questionnaire élèves.

INTRODUCTION

« Si on ne permet pas aux enfants de réfléchir sans ces outils [d'intelligence artificielle générative], ils risquent de ne plus penser par eux-mêmes, car la bouillie arrivera par la machine, avec pour conséquences une détérioration de leur libre arbitre et de leur créativité. » C'est du moins ce qu'estime Laurence Devillers (2025, p. 22), professeure en informatique appliquée aux sciences sociales et spécialiste des enjeux liés à l'IAG. Pourtant, l'externalisation de certaines opérations mentales n'est pas nouvelle : l'école l'a déjà connue, et la calculatrice à elle seule en constitue un exemple saillant. Mais avec l'avènement de l'IAG,¹ un nouveau seuil est franchi : celle-ci n'automatise plus seulement une procédure ; elle peut proposer, en lieu et place de l'élève, une structuration et un raisonnement qui semblent « prêts à l'emploi ». C'est ce glissement qui inquiète : si l'outil fournit « la pensée en kit », l'entraînement à penser par soi-même, et avec lui le jugement et la créativité, risquent bel et bien d'en pâtir.

Et la question ne porte plus uniquement sur l'usage des outils, mais sur ce que l'on observe réellement lorsque l'on évalue une production scolaire. Un exposé oral réussi donne parfois l'illusion d'une évidence : l'élève parle avec assurance, les idées s'enchaînent, le vocabulaire est pertinent, la structure est claire. Pourtant, derrière cette apparente fluidité, tout se joue dans une zone moins visible, mais décisive pour l'enseignement des langues : la compréhension réelle du sujet, l'appropriation des idées, la capacité à répondre à une question imprévue relative à l'objet présenté, tout en maintenant un discours un tant soit peu singulier et personnel. C'est précisément cette zone, habituellement inférée à partir de la performance observable, qui m'a conduit à m'intéresser à l'évaluation de l'exposé oral à l'ère de l'IAG.

En tant qu'enseignant remplaçant durant mes années universitaires, puis actuellement enseignant en formation pour le secondaire I et II, j'ai eu l'opportunité d'observer, non sans enthousiasme, la richesse de certains exposés : des élèves qui parvenaient à « tenir » un discours en anglais, à organiser leurs idées et à prendre une place devant la classe. Dans ces moments, l'exposé n'est pas une simple restitution : il devient un espace d'apprentissage où l'élève cherche ses mots, reformule, et découvre que parler une langue, c'est aussi penser avec elle.

Or, depuis quelques mois, un phénomène inédit a progressivement déplacé cet équilibre : la généralisation d'outils d'IA qui interviennent directement dans la préparation des exposés. Lors de mes stages et de remplacements, j'ai vu des élèves s'appuyer sur ces outils pour préparer des exposés avec une efficacité de production brute parfois impressionnante. Pourtant, un décalage apparaissait lors des échanges, comme si la qualité de la prestation devenait moins informative quant à la compréhension et à l'appropriation.

¹ À ce jour, en mars 2026, les principaux outils d'IAG grand public sont notamment ChatGPT (OpenAI), Gemini (Google), Copilot (Microsoft), Perplexity (Perplexity AI) et Claude (Anthropic).

Dans les classes du secondaire I — sur lesquelles je me focalise pour cette étude — ces usages semblent très hétérogènes : correction linguistique, recherche d'idées, structuration, rédaction à différents degrés. Et dès lors qu'une partie de la préparation se déroule hors de la classe, l'usage réel de l'IAG devient difficile à circonscrire : elle peut intervenir partiellement ou massivement, sans que l'enseignant puisse identifier avec précision ce qui relève du travail propre de l'élève et ce qui relève d'une production algorithmique. Dans les différents cantons romands, cette difficulté est d'autant plus marquée que, pour le secondaire I, l'usage de l'IAG ne fait pas encore l'objet de prescriptions détaillées : il dépend ainsi largement, en pratique, des consignes fixées par les enseignant·e·s.

C'est ainsi que la question devient professionnelle avant même d'être morale : comment maintenir une évaluation juste et formatrice lorsque les conditions de production se transforment aussi drastiquement ? Une évaluation n'est pas seulement un classement, elle oriente ce que l'on valorise et, indirectement, ce que les élèves apprennent à privilégier. Si une grille traditionnelle récompense surtout la qualité formelle d'un discours figé, elle risque, dans un contexte d'IAG, de surestimer des indices de performance qui ne sauraient garantir, à eux seuls, l'appropriation du contenu.

Ce constat a nourri mon intention de recherche : comprendre, dans une situation authentique de classe, ce que change l'IA dans la préparation d'un exposé oral en anglais, et ce que ces changements impliquent pour l'évaluation. Le travail présenté ici s'appuie sur un dispositif comparatif mené avec les mêmes élèves : un premier exposé préparé sans IA, puis un second exposé préparé avec un usage encadré de l'IA. L'objectif n'est pas de formuler une condamnation générale ni — bien évidemment — de célébrer l'outil, mais de décrire et d'interpréter, à partir de données concrètes, les effets observables sur la performance orale et sur le rapport des élèves à leur travail. Pour éclairer cette dimension moins visible, un questionnaire administré après chaque exposé, avant la communication des notes, permettra aussi de recueillir le point de vue des élèves sur leur compréhension, leur sentiment d'appropriation et, dans la seconde condition, sur la manière dont l'IA a été mobilisée.

À travers cette démarche, l'enjeu est double. Il s'agit d'abord de mieux comprendre si l'usage encadré de l'IA modifie surtout la mise en forme linguistique et discursive, ou s'il tend également à altérer l'appropriation du contenu. Ensuite, il sera question d'interroger, de manière pragmatique, la pertinence des critères par lesquels nous attribuons généralement une note à un exposé oral : autrement dit, de vérifier si nos outils évaluatifs restent cohérents avec ce que nous prétendons mesurer, dans un environnement d'apprentissage transformé.

1. PROBLÉMATIQUE ET CADRE THÉORIQUE

1.1. Point de départ de la réflexion : l'IA et la fragilisation de l'inférence évaluative

Dans l'évaluation d'un exposé oral, une intuition structure souvent le jugement enseignant : une prestation linguistiquement maîtrisée et solidement organisée est supposée refléter, au moins minimalement, une compréhension du sujet et une appropriation des idées. Or, l'essor des IAG modifie suffisamment les conditions de préparation pour fragiliser cette inférence.

En effet, de tels outils permettent de générer rapidement des contenus cohérents, organisés et ajustés à un niveau donné, lorsque la demande est formulée efficacement (Wang et al., 2024, p. 147). La qualité formelle d'un script préparatoire ou d'un plan peut ainsi être fortement améliorée sans que l'on puisse en déduire mécaniquement l'engagement cognitif qui a conduit à ce résultat. Le phénomène n'est pas entièrement nouveau : des exposés largement copiés (par exemple depuis Wikipédia) existent depuis environ deux décennies (Sormunen, E., & Lehtiö, L., 2011, p. 2), mais ce biais semble aujourd'hui exacerbé (Gerlich, 2025, p. 2-3) : dans une tâche d'exposé oral, l'écart potentiel entre la performance présentée et la maîtrise réelle du contenu devient une hypothèse de plus en plus plausible, et non une exception. Ces éléments conduisent aux questions suivantes : si la préparation peut être médiée par une IA, que mesurent exactement les grilles traditionnelles d'évaluation de l'oral, et dans quelle mesure leurs critères restent-ils alignés avec les compétences que l'enseignement des langues vise à développer ?

1.2. L'exposé oral en langue étrangère : une tâche complexe au service de compétences multiples

L'exposé oral en langue étrangère ne se réduit pas à « parler correctement ». La littérature le décrit comme une tâche complexe, dans laquelle l'élève doit comprendre un contenu, sélectionner des informations pertinentes, construire une progression logique et maintenir une intelligibilité suffisante pour un auditoire (Tano, 2021, p. 2-5). Il mobilise des ressources linguistiques (lexique, syntaxe, connecteurs), des compétences discursives (organisation, cohérence, clarté), mais aussi des dimensions interactionnelles et pragmatiques (regard, posture, gestion du rythme, capacité à répondre). Dans cette perspective, l'oral peut rendre visibles des compétences que l'écrit saisit moins directement, notamment en termes de présence communicationnelle et de gestion de l'interaction (Colognesi & Gagnon, 2022, p. 2).

Cette réalité didactique rend l'exposé particulièrement pertinent pour étudier l'impact de l'IA : l'outil intervient précisément dans les dimensions où les élèves ont traditionnellement besoin de médiation (organisation, reformulation, correction), tout en laissant ouvertes les questions de compréhension, d'appropriation et de capacité à soutenir un discours personnel.

1.3. IAG et « effet de surface » : vers un paroxysme de la décorrélation forme-compréhension ?

Avec l'IAG, un élève peut, plus que jamais, produire un texte « comme s'il l'avait écrit lui-même », ce qui brouille la frontière entre compétence productive authentique et soutien externe (Kovari, 2025, p. 6). Cette possibilité, déjà problématique pour l'écrit, s'étend à l'oral par un détour : l'élève peut mémoriser, paraphraser ou lire des segments fortement préparés, et ainsi présenter un discours formellement abouti. Certes, l'oral impose des contraintes incarnées (prononciation, fluidité, gestion de la situation), mais elles ne suffisent pas à garantir une appropriation conceptuelle : « pouvoir dire » ne prouve pas toujours « savoir ce qu'on dit », surtout si la préparation a été largement externalisée.

Ce phénomène peut être éclairé par la notion de *cognitive offloading*, qui désigne l'externalisation de certaines opérations cognitives vers des artefacts techniques ou des environnements informationnels. Dans le cas des IAG, cette externalisation peut concerner des opérations centrales de la production langagière — planification, formulation ou structuration —, ce qui tend à distendre le lien entre la performance observable et l'activité cognitive effectivement mobilisée par l'élève. Des travaux expérimentaux récents menés au Massachusetts Institute of Technology (MIT) suggèrent d'ailleurs que l'usage d'un assistant d'écriture fondé sur un modèle de langage peut s'accompagner d'un engagement cognitif plus faible et d'un sentiment d'appropriation amoindri du texte produit (Kosmyna et al., 2025, p. 2).

Dans ce contexte, l'évaluation risque de survaloriser un « effet de surface » : structure rhétorique soignée, lexique plus riche, correction grammaticale accrue, transitions plus fluides (Kovari, 2025, p. 5-6). Or, ces indices peuvent provenir d'un artefact algorithmique plutôt que d'un apprentissage langagier consolidé, si l'usage de l'IA n'a pas été orienté vers la compréhension et la métacognition (Noroozi et al., 2024, p. 374). La conséquence pédagogique est directe : si les critères évalués ne distinguent pas suffisamment la qualité du produit et la solidité du processus, l'évaluation peut attribuer des points à une performance dont la part réellement appropriée par l'élève demeure incertaine. Chaudhry et al. (2023, p. 23-24) concluent ainsi leur recherche : il est nécessaire de réinterroger les approches d'évaluation existantes lorsque les outils génératifs modifient la nature des productions attendues. Mais comment ? la question reste à ce stade en suspens.

1.4. Validité et équité : deux exigences mises sous tension

Dans les sciences de l'éducation, la validité d'une évaluation renvoie à la qualité des inférences que l'on tire à partir d'une performance observée, en lien avec les objectifs visés et la situation d'apprentissage (De Ketele, 2011, p. 90-91 ; Pasquini, 2019, p. 65-66). Dans une perspective de validité, la question n'est pas seulement : « la note est-elle cohérente ? », mais : « que

signifie-t-elle réellement au regard des compétences visées ? ». Dans un contexte d'IAG, cette interrogation se renforce : si l'on continue d'évaluer prioritairement la qualité formelle du discours, tout en sachant que cette qualité peut être largement assistée, l'inférence « maîtrise linguistique et discursive = compétence de l'élève » devient moins assurée.

Et cette question est indissociable de l'équité. Kane (2010, p. 178-182) rappelle que l'évaluation doit être juste et fondée sur des critères explicites, appliqués de manière appropriée aux conditions effectives de travail. Or, l'IA introduit de nouvelles sources d'inégalités : maîtrise variable du *prompting*,² capital numérique, accès à des outils ou à des pratiques de vérification. Cotton et al. (2024, p. 230) et Kovari (2025, p. 2) soulignent que l'enjeu n'est pas simplement d'interdire ou d'autoriser, mais de prévenir des effets différenciés : une interdiction peut pénaliser des élèves dépourvus d'autres soutiens, tandis qu'une autorisation sans cadre peut favoriser ceux qui savent déjà exploiter l'outil avec efficacité. Dans un système évaluatif, ces écarts ne sont pas neutres : ils peuvent se transformer en différences conséquentes de notes, donc en différences de reconnaissance scolaire.

Au secondaire I, cette dimension est particulièrement sensible, car les écarts d'autonomie et de maturité scolaire sont importants (Bélanger & Farmer, 2012, p. 174–175). Une évaluation valide et équitable doit donc rendre visible ce qui relève de compétences travaillées en classe et limiter la part d'avantage liée à une optimisation technique externe. C'est ici que l'éthique évaluative intervient : elle suppose transparence, cohérence entre consignes et critères et alignement avec les finalités éducatives (De Ketele, 2011 ; Kane, 2010 ; Pasquini, 2019), dans un contexte où les outils d'IA soulèvent également des enjeux éthiques spécifiques pour l'éducation (Akgun & Greenhow, 2022).

1.5. Vers une adaptation raisonnée des critères : justification d'une grille « IA-informée »

Si l'on admet que l'IA modifie les conditions de production, l'adaptation des instruments d'évaluation devient une hypothèse de travail légitime, et non une réaction défensive. La littérature récente converge vers quatre pistes : recentrer l'évaluation sur l'oral effectif, valoriser des traces de processus, intégrer des indices de compréhension et d'appropriation, et expliciter les attentes d'usage responsable et transparent. Une grille « IA-informée » ne vise donc pas à « punir l'IA », mais à préserver la validité de l'inférence évaluative en repondérant ce qui est aisément optimisable par assistance et en valorisant ce qui reste difficilement substituable : compréhension, appropriation, capacité à répondre, personnalisation du discours.

Cette repondération est un point central. En pratique, elle revient à attribuer moins de poids au seul « contenu brut présenté » au sens d'un produit bien rédigé et bien structuré, et davantage

² L'acte de formuler des instructions destinées à orienter la génération de contenu par un système d'IAG.

à des indices de maîtrise personnelle : réponses pertinentes aux questions, exemples personnels, capacité à expliciter et à justifier, cohérence entre le discours et ce que l'élève peut défendre oralement. La grille « IA-informée » s'inscrit ainsi dans une logique d'ajustement évaluatif cohérente avec un environnement d'apprentissage transformé, sans renoncer aux exigences linguistiques, mais en rééquilibrant ce qui est considéré comme probant pour juger de la compétence de l'élève.

1.6. Question de recherche et objectifs de l'étude

La problématique développée ci-dessus conduit à la formulation de la question de recherche suivante :

Comment repenser les critères d'évaluation traditionnels de l'exposé oral en contexte scolaire d'apprentissage de langues étrangères à l'ère de l'IA, afin de garantir une mesure équitable et pertinente des compétences réellement mobilisées par les élèves ?

À partir de cette question centrale, cette étude poursuivra les objectifs suivants :

- Analyser les limites des grilles d'évaluation actuelles face à l'usage d'outils d'IA dans la préparation des exposés oraux.
- Identifier les compétences non substituables — ou difficilement automatisables — que l'exposé oral permet de mobiliser.
- Expérimenter des critères ou des dispositifs d'évaluation alternatifs intégrant ou régulant l'usage des outils d'IA.
- Proposer, à partir de ce cas, des pistes concrètes de transformation des pratiques évaluatives, susceptibles d'éclairer d'autres contextes disciplinaires.

2. MÉTHODOLOGIE

2.1. Dispositif de recherche : les grandes lignes

Cette étude s'inscrit dans une démarche exploratoire de type étude de cas, appuyée sur une quantification descriptive et une interprétation didactique contextualisée. Le dispositif repose sur une comparaison intra-sujets : chaque élève réalise deux prestations successives, ce qui permet de limiter l'effet de certaines différences interindividuelles, telles que le niveau de langue initial ou l'aisance à l'oral, sans toutefois les neutraliser entièrement. Le premier exposé est réalisé sans recours à l'IAG, tandis que le second intègre son usage autorisé et encadré. Cette structure comparative permet d'observer des écarts de performance et de perception entre les deux conditions, sans isoler causalement l'effet propre de l'IA.

2.2. La nature des données et leurs procédures respectives de recueil

L'étude a été conduite dans deux classes du secondaire I dans le cadre d'un cours d'anglais. Le corpus se compose de 27 élèves, anonymisés dans l'analyse sous la forme d'identifiants codés (E01 à E27), conformément aux principes de confidentialité et de protection des données en recherche éducative. Les classes présentent une hétérogénéité attendue du secondaire I, notamment en termes de niveau linguistique, d'autonomie et de rapport à l'oral. Les élèves appartiennent à différentes options scolaires : options 2, 3 et 4. L'option 2 correspond au parcours pré-gymnasial ; les options 3 et 4 regroupent des élèves estimés tendanciellement plus faibles scolairement, avec un rythme d'apprentissage plus lent et/ou irrégulier.

Le dispositif pédagogique repose sur la réalisation de deux exposés oraux successifs, intégrés à la progression didactique ordinaire. Le premier exposé, réalisé sans recours à l'intelligence artificielle, consistait à présenter un métier imaginaire en 2050 (cf. annexe A). Les élèves devaient nommer le métier, décrire une journée type, présenter les qualités nécessaires et exprimer leur opinion personnelle, dans une présentation orale structurée de 3 à 4 minutes par binôme. L'usage de l'IA y était explicitement interdit, tandis que les recherches sur Internet, supervisées, ainsi que le dictionnaire restaient autorisés. La préparation s'est déroulée sur quatre périodes d'enseignement, uniquement en classe, suivies d'une cinquième période dédiée aux présentations orales.

Le second exposé, réalisé avec IA autorisée, portait sur la présentation d'un plat célèbre d'un autre pays (cf. annexe B). Les élèves devaient en décrire l'origine, les ingrédients, la préparation, le contexte de consommation, ainsi que formuler une appréciation personnelle et une comparaison culturelle. La structure de la tâche (recherche, rédaction, présentation orale) est restée volontairement comparable afin d'assurer la cohérence du dispositif. La préparation s'est effectuée uniquement en classe une nouvelle fois, cette fois sur trois périodes pour une

tâche d'ampleur similaire, suivies de la passation orale. L'usage de l'IA était autorisé dans un cadre explicite : les élèves pouvaient l'utiliser durant les leçons prévues à cet effet, en salle informatique, comme outil d'aide ; ils devaient être capables d'expliquer leur contenu, de transmettre les prompts utilisés et de réécrire leur script de manière personnelle (cf. annexe C).

Le choix d'un ordre fixe (sans IA puis avec IA) répond à une contrainte pédagogique réaliste et à une logique éthique : il aurait été didactiquement incohérent d'autoriser l'IA dans un premier temps puis de l'interdire ensuite, ou de l'autoriser d'abord à certains élèves puis aux autres, ce qui aurait créé des conditions d'apprentissage inéquitables et pédagogiquement peu justes. Toutefois, ce séquençage implique la prise en compte d'un effet d'apprentissage, les élèves étant déjà familiarisés avec la structure de l'exposé lors de la seconde tâche. Il convient également de reconnaître l'influence possible de variables interindividuelles non contrôlées, telles que l'intérêt pour le thème ou l'affinité préalable avec le sujet, susceptibles d'affecter l'engagement et la qualité de la prestation, sans être directement saisies par les enquêtes menées. Ces aspects seront explicités par la suite, dans les limites de cette étude.

Les données recueillies sont de trois natures complémentaires. Premièrement, des données de performance scolaire ont été collectées à travers les notes attribuées aux deux exposés, calculées sur 25 points selon un barème critérié. Deuxièmement, des données perceptives ont été recueillies au moyen de questionnaires élaborés spécifiquement pour cette recherche (cf. annexes D et E), à partir des dimensions retenues dans le cadre théorique³ ; ils ont été administrés aux élèves une leçon après chaque passation, avant la communication des notes, afin de limiter l'influence d'un biais de justification a posteriori. Les élèves devaient indiquer leur degré d'accord (de 1 à 10) avec différentes affirmations portant notamment sur leur investissement, leur plaisir, leur compréhension du sujet, leur aisance, leur sentiment d'appropriation du discours et leur progression à l'oral. Pour le second exposé, des items supplémentaires ont été intégrés afin d'explorer spécifiquement le rapport à l'IA, notamment liés à l'organisation du travail, à la charge cognitive, à la vérification des informations, au degré de modification personnelle du contenu et au sentiment de maîtrise du travail (cf. annexe E). Troisièmement, des traces du dispositif pédagogique (consignes, grilles d'évaluation, structure des tâches) ont été intégrées à l'analyse afin d'assurer la transparence méthodologique.

Concernant l'évaluation, deux grilles distinctes ont été mobilisées. Le premier exposé a été évalué à l'aide d'une grille standard (ou traditionnelle) de présentation orale (cf. annexe F),⁴ totalisant 25 points et fondée sur six critères : contenu présenté, organisation du discours, qualité linguistique (grammaire et vocabulaire), prononciation et fluidité, attitude de

³ Les dimensions retenues pour l'élaboration du questionnaire (sentiment de maîtrise, compréhension perçue, engagement dans la tâche, etc.) s'inspirent de travaux portant sur le sentiment d'efficacité personnelle et l'auto-régulation des apprentissages (p. ex. Bandura, 1997 ou encore Pintrich & De Groot, 1990).

⁴ La grille standard retenue ici a été élaborée à partir d'échanges avec des collègues et de la consultation de plusieurs grilles d'évaluation d'exposés oraux utilisées au sein de l'établissement scolaire de mon stage.

communication (regard, posture), ainsi que respect du temps et des consignes. Le second exposé a été évalué à l'aide d'une grille « IA-informée » (cf. annexe G), également sur 25 points, mais caractérisée par une repondération des critères : la part du contenu formel a été réduite, tandis que les dimensions d'appropriation et de compréhension ont été davantage valorisées, notamment la pertinence des réponses personnelles et la capacité à expliciter le contenu présenté. Cette adaptation repose sur l'hypothèse que l'IA améliore plus aisément la structuration et la qualité formelle du discours que l'appropriation conceptuelle et l'engagement cognitif.

2.3. Les méthodes d'analyse des données

L'analyse des données s'inscrit dans une approche principalement quantitative descriptive, complétée par une interprétation didactique contextualisée. Les données ont été reportées et structurées dans des tableurs Excel comprenant, pour chaque élève, l'identifiant anonymisé, l'option scolaire, la note obtenue et les réponses au questionnaire (Q1 à Q21), afin d'assurer la traçabilité du traitement.

Les notes des deux exposés ont été analysées à l'échelle individuelle (comparaison intra-élève) et collective (moyennes générales et moyennes par option), permettant d'identifier à la fois des tendances générales et des variations selon les profils. Les résultats scolaires ont été considérés indépendamment des questionnaires, ces derniers n'intervenant pas dans la note mais servant d'outil d'éclairage des processus sous-jacents à la performance.

Les réponses aux questionnaires, codées sur une échelle de 1 à 10, ont été traitées par calcul de moyennes par item et par condition (sans IA / avec IA). Une attention particulière a été accordée aux items directement liés à la problématique, notamment la compréhension du sujet, le sentiment d'appropriation du discours, la maîtrise de ce qui est dit durant l'exposé et la perception de progression à l'oral. Les items spécifiques à l'IA ont permis d'analyser le rôle perçu de l'outil (organisation, simplification, rédaction, vérification) ainsi que le degré d'autonomie maintenu par les élèves.

Enfin, l'interprétation s'appuie sur une triangulation des données⁵ : comparaison des notes, analyse des perceptions déclarées et prise en compte du contexte pédagogique (temps de préparation, ordre des tâches, cadre d'usage de l'IA), dans le but de renforcer la validité interne en recherchant des convergences d'indices empiriques plutôt qu'en isolant un effet causal attribuable à la seule présence de l'IA.

⁵ La triangulation consiste à croiser plusieurs sources ou types de données afin de renforcer la solidité de l'interprétation en recherchant des convergences plutôt qu'en s'appuyant sur un seul indicateur (Flick, 2004).

3. PRÉSENTATION ET ANALYSE DES RÉSULTATS

3.1. Nature des mesures et précautions analytiques

Deux types de données doivent être clairement distingués, car elles n'ont pas le même statut épistémologique. D'une part, les notes (résultat sur 6, converti depuis une note sur 25) constituent une mesure de performance issue d'une évaluation critériée réalisée par l'enseignant selon une grille explicite (cf. annexes F et G).

D'autre part, les réponses au questionnaire (cf. annexes D et E) relèvent d'une auto-évaluation par les élèves (échelle 1–10), administrée une leçon après la passation, avant la communication des notes. Ces items mesurent donc un ressenti (perception de compréhension, de maîtrise, d'appropriation, etc.) et non une indication directe de compétence. Les scores élevés à des items tels que « J'ai réellement compris le sujet » (Q3) ou « Pendant l'exposé, je savais à tout moment ce que je disais » (Q7) doivent être analysés comme des indicateurs perceptifs, à confronter aux performances évaluées.

Précisons que, dans le questionnaire IA (exposé 2), un item est rendu singulier par son format. L'item Q15 n'est pas une échelle 1–10 : il s'agit d'une question à choix, codée en variables binaires (0/1) par modalité (idées, structurer, rédiger, corriger, reformuler, comprendre, s'entraîner, autre). Les moyennes correspondantes s'interprètent alors comme des proportions de sélection (par exemple 0.63 signifie qu'environ 63% des élèves ont coché la modalité).

3.2. Évolution des performances : comparaison des notes (sans IAG vs avec IAG)

Au niveau général, la moyenne de classe passe de 4.91/6 (exposé 1 sans IA) à 4.99/6 (exposé 2 avec IA, évalué avec la grille IA-informée), soit une variation positive (+0.08) d'ampleur très modérée. Par lecture par option, on observe (cf. tableau 1) une progression marginale en O2, une stabilité en O3 et une progression plus marquée en O4, ce qui contribue à une légère hausse de la moyenne générale. Dans l'ensemble, les résultats observés restent donc globalement proches d'une condition à l'autre.

Tableau 1 — Moyennes des notes par option (avec vs sans IA)

Option	Moyenne sans IA	Moyenne avec IA	Δ
O2	5.15	5.20	+0.05
O3	4.90	4.90	+0.00
O4	4.51	4.74	+0.23

Il est utile de rappeler que l'étude vise une quantification descriptive basée sur un échantillon de données restreint, et que par conséquent ces écarts devront par la suite être interprétés comme de potentiels ordres de grandeur et non comme des effets statistiquement établis.

3.3. Effet de la grille : comparaison note IA-informée vs re-cotation avec la grille standard

Un point de comparaison méthodologiquement informatif consiste à estimer les notes qu'aurait produit, pour l'exposé 2, l'application de la grille standard utilisée pour l'exposé 1 (cf. annexe F). Dans ce cadre hypothétique, la moyenne générale atteindrait 5.21/6, alors que la moyenne effectivement obtenue pour l'exposé 2 avec la grille IA-informée (cf. annexe G) est de 4.99/6. L'écart est donc de -0.22 point.

Cet écart apparaît aussi par option :

Tableau 2 — Moyennes des notes par option (grille IA-informée vs grille standard simulée)

Option	Grille IA-informée	Grille standard simulée	Δ
O2	5.20	5.38	+0.18
O3	4.90	5.15	+0.25
O4	4.74	5.00	+0.26

Autrement dit, à prestation constante (exposé 2), la grille standard tend à attribuer davantage de points que la grille IA-informée, quelle que soit l'option considérée. À ce stade, il ne s'agit pas de trancher la « pertinence » de l'une ou l'autre grille, mais d'établir un fait méthodologique : le choix de la grille infléchit systématiquement le niveau des résultats.

Cet effet est cohérent avec la logique de construction de la grille IA-informée. Dans la grille standard de l'exposé 1, le critère « Contenu présenté » (pertinence, clarté, exemples) pouvait rapporter jusqu'à 6 points sur 25. La grille IA-informée redistribue partiellement cette pondération en réduisant le poids du contenu formel au profit d'indicateurs censés attester une compréhension effective et une appropriation du discours, notamment la capacité à expliciter, à personnaliser et à répondre. Concrètement, le critère « Contenu présenté » est plafonné à 4 points dans la grille revisitée.

Les données issues des auto-évaluations communes (Q1–Q11) apportent un éclairage compatible avec ce déplacement de pondération. Sur l'ensemble du corpus, trois items directement liés à l'appropriation perçue évoluent de manière globalement défavorable entre les deux exposés : Q3 « j'ai réellement compris le sujet » passe de 8.66 à environ 8.00 (-0.66), Q7 « je savais ce que je disais et pourquoi » de 7.55 à 7.26 (-0.29) et Q8 « le discours était vraiment le mien » de 7.96 à 7.22 (-0.74). En agrégeant ces trois items comme indicateurs de l'appropriation perçue, la moyenne composite passe de 8.06 à 7.49, soit une baisse moyenne de 0.57 point sur 10. Cette baisse agrégée reste modérée, mais sa distribution détaillée n'est pas parfaitement uniforme : elle est nette pour Q3 et Q8, tandis que Q7 présente une légère hausse en O2 et des baisses en O3 et O4.

Dans le même temps, la comparaison des performances globales (cf. section 3.2) ne suggère pas une dégradation du niveau formel : les notes restent stables, voire légèrement en hausse selon l’option. Dans ce contexte, l’écart observé entre la re-cotation standard et la cotation IA-informée s’interprète de manière parcimonieuse comme un effet de repondération : une grille qui valorise davantage l’appropriation et la compréhension tend mécaniquement à attribuer moins de points lorsque ces dimensions sont moins présentes, moins visibles, ou moins revendiquées par les élèves, même si le produit oral demeure structuré et linguistiquement correct.

Les items spécifiques à l’IAG (Q12–Q21) vont dans le même sens, en décrivant un usage principalement tourné vers la production et la structuration plutôt que vers un approfondissement actif. Les élèves déclarent un accord élevé avec « avec l’IA, j’ai dû moins réfléchir » (≈ 7.85), tandis que le degré de modification personnelle du contenu (Q17 ≈ 4.59) et la vérification (Q19 ≈ 4.55) restent modérés. Ces résultats suggèrent que l’IA a surtout soutenu l’idéeation, la formulation et l’organisation, sans garantir, dans les pratiques déclarées, une appropriation plus profonde du contenu.

La cohérence d’ensemble entre trois éléments descriptifs — effet systématique de la re-cotation (grille standard vs IA-informée), baisse modérée mais régulière des indicateurs d’appropriation perçue, et écarts globalement modestes des résultats observés — conduit à formuler l’hypothèse suivante : dans ce dispositif, la dimension « appropriation et compréhension » constitue une piste interprétative plausible de la perte relative de points avec la grille IA-informée. Autrement dit, ce n’est pas la qualité apparente du discours qui diminue, mais la part de maîtrise personnelle que l’évaluation cherche à inférer et à valoriser, ce qui rejoint l’hypothèse centrale de l’étude : l’IAG pourrait renforcer la qualité de surface du produit oral, tout en rendant moins saillante l’appropriation cognitive effective du contenu.

3.4. Analyse des auto-évaluations individuelles (Q1–Q11) : évolutions générales

Les items Q1 à Q11 sont identiques entre les deux exposés et permettent de comparer les perceptions des élèves dans les deux conditions.

Globalement, on observe des mouvements contrastés :

Tableau 3 — Comparaison des moyennes des items selon la condition de préparation (sans IA vs avec IA)

Question	Moyenne sans IA	Moyenne avec IA	Δ	Tendance observée
Q1 « meilleur de moi-même »	7.41	7.59	+0.18	très faible hausse
Q2 « plaisir »	7.30	7.11	−0.19	très faible baisse
Q3 « j’ai réellement compris le sujet »	8.67	8.00	−0.67	baisse notable

Question	Moyenne sans IA	Moyenne avec IA	Δ	Tendance observée
Q4 « la recherche m’a aidé à construire mes idées »	6.96	6.19	-0.77	baisse notable
Q5 « à l’aise avec les mots et idées préparés »	7.15	7.11	-0.04	stabilité
Q6 « mes notes m’ont aidé sans m’empêcher de parler fluide »	6.93	7.11	+0.18	hausse très faible
Q7 « je savais ce que je disais et pourquoi »	7.56	7.26	-0.30	baisse légère
Q8 « le discours était vraiment le mien »	7.96	7.22	-0.74	baisse notable
Q9 « j’ai appris de nouvelles choses sur le sujet »	7.00	6.37	-0.63	baisse notable
Q10 « l’activité évalue des compétences importantes »	7.59	6.96	-0.63	baisse notable
Q11 « j’ai progressé à l’oral devant les autres »	6.85	7.48	+0.63	hausse notable

Sur le plan strictement analytique, deux constats ressortent. Premièrement, les perceptions liées à la compréhension et à la construction d’idées personnelles (Q3, Q4), ainsi qu’au sentiment d’appropriation (Q8), diminuent entre les deux exposés, tandis que la perception de progression orale (Q11) augmente nettement. Deuxièmement, la majorité des moyennes reste néanmoins dans une zone relativement élevée (souvent au-dessus de 6.5), ce qui indique des perceptions dans l’ensemble positives, malgré les variations.

3.5. Analyse par option sur les items les plus liés à la problématique (Q3, Q7, Q8, Q11)

Une lecture plus fine, par option, permet de préciser ces évolutions perceptives.

Pour Q3 (compréhension réelle du sujet), la baisse générale est portée principalement par l’O3 :

Tableau 4 — Compréhension perçue du sujet par option (Q3 : sans IA vs avec IA)

Option	Moyenne sans IA	Moyenne avec IA	Δ	Tendance observée
O2	8.92	8.58	-0.34	baisse légère
O3	8.88	7.25	-1.63	baisse forte
O4	8.00	7.86	-0.14	baisse légère

Pour Q7 (maîtrise et justification du discours), les mouvements divergent :

Tableau 5 — Maîtrise perçue du discours par option (Q7 : sans IA vs avec IA)

Option	Moyenne sans IA	Moyenne avec IA	Δ	Tendance observée
O2	8.17	8.25	+0.08	hausse très légère
O3	7.13	6.63	-0.50	baisse
O4	7.00	6.29	-0.71	baisse

Pour Q8 (discours perçu comme « vraiment le mien »), une baisse est observée dans les trois options, particulièrement marquée en O2 et O4 :

Tableau 6 — Caractère personnel du discours par option (Q8 : sans IA vs avec IA)

Option	Moyenne sans IA	Moyenne avec IA	Δ	Tendance observée
O2	8.42	7.50	-0.92	baisse notable
O3	7.88	7.63	-0.25	baisse légère
O4	7.29	6.29	-1.00	baisse notable

Enfin, pour Q11 (progression orale perçue), la hausse est nette dans toutes les options, particulièrement en O4 :

Tableau 7 — Progression orale perçue par option (Q11 : sans IA vs avec IA)

Option	Moyenne sans IA	Moyenne avec IA	Δ	Tendance observée
O2	7.50	7.92	+0.42	hausse
O3	6.75	7.25	+0.50	hausse
O4	5.86	7.00	+1.14	hausse forte

Ces résultats montrent un profil analytique intéressant : la progression perçue à l’oral augmente partout, tandis que l’appropriation perçue (Q8) diminue partout, et que la baisse de compréhension perçue (Q3) est particulièrement concentrée sur l’O3.

3.6. Analyse des items spécifiques à l’IA (Q12–Q21) : usage, vécu, délégation

Les items Q12–Q21 (cf. annexe E) renseignent le vécu et les pratiques en lien avec l’IA.

Au niveau général :

Tableau 8 — Moyennes des items spécifiques à l’IA (Q12 à Q21)

Question	Moyenne	Tendance observée
Q12 « j’utilisais déjà souvent ChatGPT / Gemini ... »	7.41	usage relativement élevé
Q13 « plus facile de m’organiser avec l’IA »	6.96	aide organisationnelle perçue
Q14 « avec l’IA, j’ai dû moins réfléchir »	7.85	accord élevé

Question	Moyenne	Tendance observée
Q16 « sans IA, mon exposé aurait été presque pareil »	6.15	accord moyen, influence perçue variable de l'IA
Q17 « j'ai modifié une grande partie du contenu IA »	4.59	niveau modéré
Q18 « j'ai remis en question l'IA »	4.56	niveau modéré
Q19 « j'ai vérifié certaines informations »	4.56	niveau modéré
Q20 « c'était stressant »	2.44	très faible stress déclaré
Q21 « je suis resté maître de mon travail »	6.70	agentivité perçue modérée à haute

Note : l'item Q15 est traité dans une section spécifique : 4.2. Dissociation entre performance évaluée et appropriation déclarée.

La lecture par option révèle des différences structurées. L'O2 rapporte un stress très faible (Q20 : 1.75) et un sentiment de maîtrise élevé (Q21 : 7.08), mais une auto-déclaration faible de modification substantielle (Q17 : 3.50). L'O3 combine un haut niveau « moins réfléchir » (Q14 : 8.00) et un Q21 élevé (7.13), avec des niveaux plus élevés de modification (Q17 : 5.25) et de vérification (Q19 : 5.00). L'O4 présente une organisation perçue très élevée (Q13 : 8.00), mais un sentiment de maîtrise nettement plus bas (Q21 : 5.57), avec un Q17 élevé (5.71) et un Q16 plus bas (5.14), ce qui suggère une influence perçue plus forte de l'IAG sur le produit final. Ici encore, ce sont des configurations qui sont décrites, sans qu'un mécanisme en soit déduit.

3.7. Notes vs perceptions : convergences et écarts attendus

L'analyse conjointe des notes et des questionnaires doit intégrer un principe : une convergence parfaite entre performance évaluée et auto-perception n'est pas attendue. Toujours est-il que dans les données présentes, on observe des écarts de notes globalement modestes, avec une légère hausse moyenne surtout portée par l'O4, alors que certains indicateurs perceptifs diminuent (Q3, Q4, Q8) tandis que d'autres augmentent (Q11). Sur le plan analytique, cela signifie que l'évolution des performances et l'évolution du ressenti ne se superposent pas mécaniquement. La valeur du questionnaire réside précisément dans cette complémentarité : il renseigne le vécu subjectif de la tâche (compréhension ressentie, appropriation ressentie, agentivité ressentie), que la note ne capture pas directement.

4. INTERPRÉTATION ET DISCUSSION

L'idée est à présent de mieux comprendre ce que les résultats obtenus peuvent révéler des transformations du processus d'apprentissage, de la validité de l'évaluation et du rôle de l'enseignant dans un contexte scolaire marqué par l'émergence des outils génératifs. Les interprétations proposées doivent donc être lues comme des hypothèses appuyées sur des régularités observées dans le cas étudié, et non comme des mécanismes démontrés.

4.1. Stabilité des performances et robustesse relative de la tâche d'exposé oral

Les résultats mettent d'abord en évidence des écarts de notes globalement modestes entre l'exposé sans IAG et l'exposé avec IA, malgré une réduction du temps de préparation (trois leçons au lieu de quatre, en raison d'un avancement plus rapide des élèves) et une modification des conditions de travail. Ces écarts ne sauraient être interprétés ni comme l'absence d'effet de l'IA, ni comme la preuve d'un gain substantiel d'apprentissage. Ils suggèrent plutôt que, dans ce contexte précis, la tâche d'exposé oral conserve une certaine robustesse face à l'introduction d'un outil technologique.

Cette robustesse peut s'expliquer par la nature même de l'oral évalué. L'exposé mobilise en effet des compétences difficilement externalisables : gestion de la voix, présence, posture, régulation du rythme, interaction avec l'auditoire, mobilisation en temps réel de ressources linguistiques, ainsi que capacité à reformuler (Tano, 2021). Contrairement à certaines productions écrites, la performance demeure incarnée, située et soumise à des contraintes de temporalité et d'exposition sociale qui limitent la substituabilité par un outil génératif.

Dans le contexte du secondaire I, cette dimension est particulièrement déterminante : l'élève ne peut déléguer ni la prise de parole elle-même ni l'engagement corporel qu'elle implique. L'IAG peut dès lors soutenir la préparation, mais elle ne se substitue pas aux composantes interactionnelles et performatives qui structurent l'évaluation de l'oral. Cette distinction rejoint certaines analyses critiques selon lesquelles les modèles de langage peuvent produire une cohérence discursive sans pour autant garantir une compréhension vécue du sens, ce qui sépare nettement la production algorithmique d'une activité langagière effectivement assumée par un sujet parlant (Kovari, 2025 ; Noroozi et al., 2024).

4.2. Écart entre performance évaluée et appropriation déclarée : un résultat central

L'un des résultats centraux de l'étude est la dissociation entre la performance évaluée et plusieurs dimensions auto-déclarées, en particulier la compréhension du sujet, la construction d'idées personnelles et le sentiment d'appropriation du discours. Les questionnaires mettent ainsi en évidence un vécu d'apprentissage plus contrasté que ne le laisse apparaître la seule

mesure de performance : la qualité observable du produit oral ne suffit pas à rendre compte de la manière dont les élèves disent s'être engagés cognitivement dans la tâche.

Cette dissociation s'éclaire à partir des usages déclarés : à l'item Q15 (« Pour ce travail, l'IA m'a surtout aidé à... »), les élèves décrivent un recours majoritairement orienté vers des opérations de production (génération d'idées, rédaction, mise en forme, organisation), plutôt que vers des démarches explicitement centrées sur la compréhension, la vérification ou l'appropriation conceptuelle. Cela renforce ainsi l'idée d'un « effet de surface » déjà discutée dans la littérature sur les usages éducatifs de l'IAG (Noroozi et al., 2024 ; Kovari, 2025).

Dans le cas étudié, les usages déclarés par les élèves peuvent ainsi être interprétés à la lumière des travaux sur l'externalisation cognitive liée aux technologies numériques (Gerlich, 2025 ; Kosmyna et al., 2025). Lorsque l'IA intervient principalement dans la génération d'idées, la formulation ou l'organisation du discours, une partie des opérations les plus exigeantes de la préparation semble se déplacer vers l'outil. Les données recueillies suggèrent alors moins une disparition de l'apprentissage qu'une reconfiguration des opérations mobilisées, susceptible d'éclairer le décalage observé entre performance évaluée et appropriation déclarée.

4.3. L'IAG comme outil d'efficacité et d'allègement de la charge cognitive

Les items spécifiques à l'IAG indiquent que l'outil est principalement mobilisé comme support d'efficacité dans la réalisation de la tâche : organisation facilitée, production accélérée et effort cognitif perçu comme moindre. Cette configuration est compatible avec l'hypothèse d'un allègement de la charge cognitive portant surtout sur la préparation discursive (recherche d'idées, structuration, reformulation), davantage que sur une consolidation explicite de la compréhension (Gerlich, 2025 ; Kosmyna et al., 2025).

Deux implications peuvent en être tirées. Premièrement, les capacités discursives des systèmes d'IAG dans la génération de formulations cohérentes peuvent renforcer la qualité formelle du produit oral, sans que cette élaboration constitue, en elle-même, un indice fiable de profondeur conceptuelle ou de compréhension effective. Deuxièmement, la réduction du temps de préparation observée dans le dispositif (trois périodes au lieu de quatre) s'interprète de manière parcimonieuse comme un effet d'optimisation du travail : l'outil permet de réaliser plus rapidement certaines opérations, ce qui modifie la distribution de l'effort sans supprimer toute activité cognitive. L'enjeu didactique n'est donc pas seulement l'efficacité obtenue, mais la nature des opérations intellectuelles que l'on souhaite rendre nécessaires, visibles et évaluables (Chaudhry et al., 2023).

4.4. Auto-évaluation, sentiment de maîtrise et potentielle illusion d'efficacité

Au-delà de l'appropriation déclarée, les questionnaires suggèrent également un décalage possible entre sentiment d'efficacité et apprentissage effectif. Le caractère auto-déclaratif des questionnaires constitue un élément déterminant pour l'interprétation. Les élèves estiment des niveaux significatifs de progression orale, de maîtrise et d'apprentissage, mais ces perceptions ne correspondent pas strictement aux évolutions des performances mesurées. Cette configuration met en évidence une possible dissociation entre sentiment d'efficacité et apprentissage effectif (Bandura, 1997 ; Pintrich & De Groot, 1990).

La fluidité et la cohérence des productions assistées sont susceptibles de renforcer le sentiment de compétence, en produisant une forme d'« illusion d'efficacité » liée à l'élaboration formelle du produit final (Bandura, 1997). Cet effet peut être plus marqué au secondaire I, où les capacités métacognitives et critiques sont encore en fort développement.

Les résultats empiriques s'inscrivent dans cette logique : l'IAG semble favoriser un sentiment de contrôle du travail et une réduction du stress, tout en laissant apparaître une appropriation plus nuancée du contenu. L'outil contribue ainsi à sécuriser la production, sans garantir une transformation proportionnelle de la compréhension.

4.5. L'effet de la grille IA-informée et la question de la validité évaluative

La différence observée entre la grille standard et la grille IA-informée constitue un élément clé de la discussion. La grille traditionnelle, davantage centrée sur le produit final et la qualité formelle du discours, tend à produire des résultats légèrement plus élevés pour l'exposé avec IA. À l'inverse, la grille IA-informée, qui répond aux critères en intégrant l'appropriation, les réponses personnelles et le processus, introduit une exigence plus alignée avec les conditions contemporaines de production scolaire.

Cette observation renvoie directement à la notion de validité de l'évaluation (De Ketele, 2011 ; Kane, 2010 ; Pasquini, 2019). Lorsque les conditions de production évoluent sous l'effet de technologies génératives, une évaluation exclusivement centrée sur le produit final risque de mesurer partiellement la performance de l'outil plutôt que celle de l'élève. L'introduction de l'IAG semble moins transformer immédiatement les apprentissages qu'elle ne met sous tension les méthodes d'évaluation centrées sur le produit, dès lors que les conditions de production ont changé. Cela rejoint la fragilisation de l'inférence évaluative discutée en section 1.1 (De Ketele, 2011 ; Kane, 2010).

Ainsi, dans ce contexte précis, l'adaptation de la grille d'évaluation peut être comprise non comme une concession à la technologie, mais comme une tentative cohérente de réajuster l'alignement entre objectifs d'apprentissage, pratiques pédagogiques et critères de notation.

4.6. Transformation du rapport à l'effort, au temps et au sens de l'apprentissage

Les résultats mettent en évidence une évolution du rapport des élèves au travail scolaire. La réduction du temps de préparation, conjuguée à des résultats observés globalement proches entre les deux conditions, suggère que l'IAG modifie la temporalité de l'apprentissage sans transformer nettement, à court terme, les résultats observables.

Cette optimisation technologique peut favoriser une focalisation accrue sur l'efficacité productive, au risque de déplacer l'apprentissage vers une logique de rendement et de performance rapide plutôt que vers un processus d'élaboration progressive du sens (Gerlich, 2025 ; Noroozi et al., 2024). Dans le cas étudié, les données vont dans ce sens : les élèves préparent plus rapidement, produisent des exposés aux résultats comparables, tout en déclarant un vécu d'apprentissage plus nuancé.

Cette évolution ne doit pas nécessairement être interprétée comme une perte drastique de sens, mais comme une transformation des modalités d'engagement cognitif. L'élève n'apprend pas nécessairement moins, mais autrement, dans un environnement où certaines opérations intellectuelles sont partiellement prises en charge par l'outil (Kosmyrna et al., 2025).

4.7. Enjeux éthiques et professionnels pour l'enseignement au secondaire I

Enfin, l'ensemble des résultats met en évidence des implications directes pour la pratique enseignante. L'IAG n'apparaît ni comme un substitut total à l'apprentissage, ni comme un simple outil neutre, mais comme un médiateur pédagogique qui reconfigure les interactions entre élève, savoir et évaluation.

Les résultats de cette étude soulignent alors une « double contrainte » éducative face à l'IAG : l'interdiction totale ignore les transformations sociétales, tandis qu'un usage non encadré risque de fragiliser l'appropriation des savoirs et l'esprit critique (Cotton et al., 2024 ; Kovari, 2025). Dans ce contexte, le rôle de l'enseignant évolue : il ne s'agit plus seulement d'évaluer une production finale, mais d'accompagner les élèves dans un usage réflexif des outils numériques, en valorisant les dimensions non automatisables de l'apprentissage, telles que la compréhension, l'appropriation et la capacité à justifier son discours. Ce point rejoint aussi les enjeux d'équité discutés plus haut (Kane, 2010 ; Akgun & Greenhow, 2022).

4.8. Ouverture réflexive : agentivité, jugement et finalités éducatives à l'ère de l'IA

Au-delà des résultats strictement empiriques, cette recherche invite à une réflexion plus large sur les finalités de l'éducation dans un contexte technologique en mutation. L'IAG ne remet pas seulement en question les modalités de production des travaux scolaires, mais interroge plus profondément le sens de l'évaluation et la place de l'élève comme sujet apprenant.

En effet, le risque majeur n'est pas que la machine remplace l'humain, mais que l'activité intellectuelle soit progressivement déléguée à des systèmes techniques, au point de fonctionner comme une sorte de prothèse cognitive, transformant le rapport à l'effort, au jugement et à la construction du savoir. Dans ce cadre, l'enjeu éducatif central devient la préservation de l'agentivité de l'élève, c'est-à-dire sa capacité à s'approprier un discours, à comprendre un contenu et à répondre personnellement de ses apprentissages (Bélanger & Farmer, 2012).

L'automatisation croissante des productions intellectuelles pourrait conduire l'éducation à redéfinir ce qu'elle cherche réellement à observer lors de présentations orales : non plus principalement la capacité à produire un discours correct, mais la faculté de juger, de comprendre et de donner du sens aux savoirs mobilisés. Dans cette perspective, les résultats de l'étude peuvent être interprétés comme révélateurs d'une mutation plus large : l'IAG n'annule pas l'apprentissage, mais elle en déplace les contours. Elle agit comme un outil de médiation discursive et d'optimisation du processus de production, tout en amenant l'institution scolaire à réinterroger la validité de ses évaluations, le sens de ses exigences et la place centrale de l'élève comme acteur réflexif de son propre apprentissage.

CONCLUSION

Cette recherche a examiné, dans un contexte authentique d'enseignement de l'anglais au secondaire I, la pertinence des critères traditionnels d'évaluation de l'exposé oral à l'ère de l'intelligence artificielle générative. À partir d'un dispositif comparatif simple — un exposé préparé sans IA, puis un exposé de même ampleur préparé avec un usage autorisé et encadré — l'objectif n'était ni de condamner ni de promouvoir l'outil, mais d'analyser, à partir de données empiriques (notes, auto-évaluations et usages déclarés), les déplacements qu'il introduit dans la préparation, la performance et l'inférence évaluative associées à une tâche orale en langue étrangère.

Les résultats mettent d'abord en évidence une relative stabilité des performances évaluées entre les deux conditions, malgré une réduction du temps de préparation et des modalités de travail différentes. Les moyennes de classe demeurent proches, indépendamment de l'option. Ce constat ne permet pas de conclure à une absence d'effet de l'IAG, mais invite à nuancer l'idée d'un remplacement immédiat des compétences par l'outil. Dans une tâche orale située, l'élève reste l'agent effectif du discours et mobilise des dimensions incarnées — prise de parole, présence, gestion du temps, interaction — qui ne peuvent être entièrement déléguées à un système génératif.

Cette stabilité du produit évalué ne signifie toutefois pas une invariance du processus d'apprentissage. La mise en relation des notes et des auto-évaluations fait apparaître une dissociation, cohérente bien que modérée : les élèves déclarent une progression à l'oral et une organisation facilitée, tout en rapportant un sentiment d'appropriation et de construction personnelle des idées plus nuancé. Pris ensemble, ces indices suggèrent que l'IAG agit principalement comme un outil de médiation de la production discursive : elle facilite l'idéation, la structuration et la reformulation, sans garantir une appropriation conceptuelle équivalente. Les usages déclarés confirment cette tendance, en renvoyant majoritairement à des fonctions de production et d'organisation du discours ; dans ce cadre, la réduction du temps de préparation observée peut être interprétée comme un effet d'optimisation du travail, qui redistribue l'effort sans supprimer toute activité cognitive.

Un résultat particulièrement structurant de cette étude réside dans la comparaison entre la grille d'évaluation standard et la grille dite « IA-informée ». La re-cotation simulée suggère que la grille traditionnelle, centrée sur le produit final (contenu, organisation, correction linguistique), tend à attribuer des notes légèrement plus élevées à une prestation préparée avec l'IAG que la grille adaptée, qui pondère davantage l'appropriation, les réponses personnelles et le processus. Cette différence, même modérée, rappelle que l'évaluation est une opération d'inférence : lorsqu'un outil peut améliorer la qualité de surface d'un discours, une grille trop centrée sur la forme risque de mesurer partiellement une performance assistée plutôt que les compétences

effectivement mobilisées. Inversement, une grille IA-informée apparaît, dans ce contexte, comme une tentative cohérente de restaurer l'alignement entre objectifs d'apprentissage, conditions de production et critères évalués.

Ces résultats permettent d'apporter une réponse prudente mais opérationnelle à la question de recherche : repenser les critères d'évaluation traditionnels de l'exposé oral à l'ère de l'IAG consiste, dans une large mesure, à déplacer le centre de gravité du jugement du seul produit discursif vers des indicateurs mieux liés à l'appropriation et à l'agentivité de l'élève. Concrètement, cela implique (a) de réduire la dépendance de la note à la seule qualité formelle du discours, devenue plus facilement optimisable, (b) de pondérer davantage la compréhension attestée en situation, la personnalisation, la capacité à expliciter et à répondre, et (c) d'intégrer des éléments de processus permettant de rendre visible la part de travail réellement assumée par l'élève. Dans l'ensemble, l'étude répond aux objectifs annoncés : elle suggère les limites de grilles centrées principalement sur le produit dans un contexte d'usage de l'IAG, met en lumière certaines dimensions de la performance orale plus difficilement automatisables, explore une grille IA-informée intégrant des critères d'appropriation et de processus, et dégage des pistes de transformation des pratiques évaluatives.

Plusieurs limites doivent toutefois être reconnues. L'étude repose sur un échantillon restreint de 27 élèves, ce qui limite la portée généralisable des résultats. Le choix d'un ordre fixe (sans IAG puis avec IAG), certes cohérent sur le plan pédagogique, introduit un effet d'apprentissage inévitable, les élèves étant familiarisés avec la structure de l'exposé lors de la seconde passation. Les thématiques des exposés, bien que comparables, pouvaient quant à elles susciter des degrés variables d'intérêt ou d'affinité. Les questionnaires reposent sur des auto-évaluations, qui renseignent un vécu plutôt qu'une mesure directe d'apprentissage. La procédure de cotation constitue également une limite : les deux exposés ont été évalués par le même enseignant, ce qui assure une cohérence intra-classe mais ne permet pas d'estimer la fidélité inter-juges. Enfin, l'absence de groupe contrôle parallèle, de *randomisation* et de suivi longitudinal empêche d'isoler causalement l'effet propre de l'IAG.

Ces limites ouvrent des perspectives de prolongement. Des études ultérieures pourraient élargir l'échantillon à plusieurs classes et niveaux, introduire des groupes parallèles (usage autorisé vs usage régulé, ou IAG encadrée vs IAG libre), et mobiliser des données qualitatives afin de documenter plus finement les stratégies cognitives. Il conviendrait notamment de creuser les raisons pour lesquelles l'IAG semble fragiliser l'appropriation de manière hétérogène selon les profils d'élèves, comme le suggère la baisse de compréhension plus marquée observée ici chez les élèves de l'option 3. De plus, le développement de dispositifs métacognitifs permettrait d'aider les élèves à prendre conscience d'une potentielle « illusion d'efficacité », où la qualité formelle du produit final occulte un appauvrissement du processus de réflexion sous-jacent. Enfin, l'utilisation de traces de prompts, de versions successives ou de justifications de choix permettrait de mieux saisir la part de travail réellement assumée par l'élève.

Au-delà de ses résultats empiriques, cette étude invite à une réflexion plus large sur les finalités éducatives à l'ère des technologies génératives. L'IAG n'oblige pas seulement à adapter des outils ; elle rend plus visible une distinction souvent implicite : celle entre performance de surface et appropriation effective. Lorsque la forme devient plus aisément optimisable, l'école est conduite à expliciter ce qu'elle valorise réellement.

En ce sens, repenser l'évaluation de l'exposé oral à l'ère de l'IAG ne relève pas uniquement d'un ajustement technique, mais d'un effort de cohérence : rendre l'évaluation plus valide et plus équitable, tout en préservant l'exigence éducative fondamentale consistant à former des élèves capables de raisonnement et de pensée autonome. En définitive, l'enjeu de l'école de demain ne se situe pas dans la rivalité avec des technologies aveugles, mais dans sa capacité à demeurer ce lieu où l'on apprend à ne pas se soumettre aux seules logiques de l'efficacité, car former des esprits critiques, c'est les préparer à affronter des futurs incertains.

BIBLIOGRAPHIE

Akgun, S., & Greenhow, C. (2022). Artificial intelligence in education: Addressing ethical challenges in K-12 settings. *AI and Ethics*, 2(3), 431–440. <https://doi.org/10.1007/s43681-021-00096-7>.

Bandura, A. (1997). *Self-efficacy : The exercise of control*. New York, NY : W. H. Freeman.

Bélanger, N., & Farmer, D. (2012). Autonomie de l'élève et construction de situations scolaires. Études de cas à l'école de langue française en Ontario (Canada). *Éducation et Sociétés*, 29(1), 173–191. <https://doi.org/10.3917/es.029.0173>.

Chaudhry, I. S., Sarwary, S. A. M., El Refae, G. A., & Chabchoub, H. (2023). Time to revisit existing student's performance evaluation approach in higher education sector in a new era of ChatGPT — A case study. *Cogent Education*, 10(1), Article 2210461. <https://doi.org/10.1080/2331186X.2023.2210461>.

Colognesi, S., & Gagnon, R. (2022). Éditorial : Évaluer les performances orales sans les dénaturer ? *Évaluer – Journal international de recherche en éducation et formation*, 8(2), 1–4. <https://doi.org/10.48782/e-jiref-8-2-1>.

Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, 61(2), 228–239. <https://doi.org/10.1080/14703297.2023.2190148>.

De Ketele, J.-M. (2011). L'évaluation et le curriculum : les fondements conceptuels, les débats, les enjeux. *Les Dossiers des sciences de l'éducation*, 25, 89–106. <https://doi.org/10.4000/dse.1022>.

Devillers, L. (2025). *L'IA, ange ou démon ? Le nouveau monde de l'invisible*. Éditions du Cerf.

Flick, U. (2004). Triangulation in qualitative research. In U. Flick, E. von Kardorff, & I. Steinke (Eds.), *A companion to qualitative research* (p. 178-183). London, England : Sage.

Genelot, S. (2023). Quelles sont les pratiques et les représentations de l'évaluation chez les enseignants ? In Cnesco (Éd.), *Conférence de consensus du Cnesco « l'évaluation en classe, au service de l'apprentissage des élèves : notes des experts »* (p. 46-61). Cnesco-Cnam.

Gerlich, M. (2025). AI tools in society: Impacts on cognitive offloading and the future of critical thinking. *Societies*, 15(1). <https://doi.org/10.3390/soc15010006>

Kane, M. T. (2010). Validity and fairness. *Language Testing*, 27(2), 177–182. <https://doi.org/10.1177/0265532209349467>.

Kosmyna, N., Hauptmann, E., Yuan, Y. T., Situ, J., Liao, X.-H., Beresnitzky, A. V., Braunstein, I., & Maes, P. (2025). *Your brain on ChatGPT : accumulation of cognitive debt when using an AI assistant for essay writing tasks*. MIT Media Lab.

Kovari, A. (2025). Ethical use of ChatGPT in education: Best practices to combat AI-induced plagiarism. *Frontiers in Education*, 9, Article 1465703. <https://doi.org/10.3389/feduc.2024.1465703>.

Noroozi, O., Soleimani, S., Farrokhnia, M., & Banihashem, S. K. (2024). Generative AI in education: Pedagogical, theoretical, and methodological perspectives. *International Journal of Technology in Education*, 7(3), 373-385. <https://doi.org/10.46328/ijte.845>

Pasquini, R. (2019). Élargir conceptuellement le modèle de l'alignement curriculaire pour comprendre la cohérence des pratiques évaluatives sommatives notées des enseignants : enjeux et perspectives. *Mesure et évaluation en éducation*, 42(1), 63-92. <https://doi.org/10.7202/1066598ar>.

Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, 82(1), 33–40. <https://doi.org/10.1037/0022-0663.82.1.33>.

Sormunen, E., & Lehtiö, L. (2011). *Authoring Wikipedia articles as an information literacy assignment: Copy-pasting or expressing new understanding in one's own words?* *Information Research*, 16(4).

Tano, M. (2021). L'exposé oral en langue étrangère: une tâche complexe au service du développement des compétences. *Recherches en didactique des langues et des cultures. Les cahiers de l'Acedle*, 18(18-1). <https://doi.org/10.4000/rdlc.8714>.

Wang, N., Wang, X., & Su, Y.-S. (2024). Critical analysis of the technological affordances, challenges and future directions of generative AI in education: A systematic review. *Asia Pacific Journal of Education*, 44(1), 139–155. <https://doi.org/10.1080/02188791.2024.2305156>.

ANNEXES

Annexe A — Consignes données aux élèves (exposé 1 sans IA)

🧑 Let's imagine it's 2050. What jobs are humans doing? 🧑



1. Consigne :

En binôme, vous présentez un métier (réel ou inventé) de 2050. Parlez comme si vous étiez vraiment en 2050. Votre présentation doit être claire et bien organisée.

2. Ce que vous devez dire :

- Donner le nom du métier (si c'est un métier imaginaire, vous pouvez l'inventer).
- Décrire une journée type pour ce métier (utilisez au moins 3 phrases avec *have to / don't have to* pour les obligations).
- Les qualités nécessaires pour exercer ce métier.
- Si vous aimeriez faire ce métier, et pourquoi.

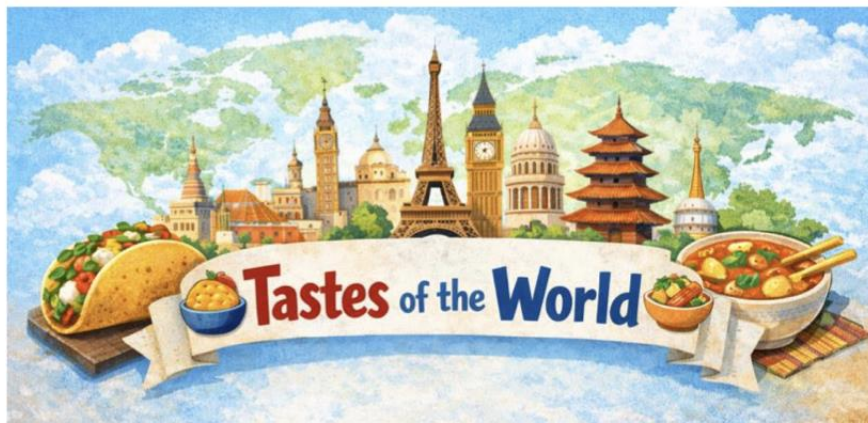
3. Étapes à suivre :

- 1) Imaginez que vous êtes en 2050. Choisissez un métier qui existe ou pourrait exister.
- 2) Rédigez le texte de l'exposé selon un plan simple : introduction – présentation du métier (👉 voir partie 2 : Ce que vous devez dire 👉) – conclusion.
- 3) Présentez le métier à l'oral (notes autorisées, mais pas le script complet).

4. Règles importantes :

- Recherches internet et dictionnaire : autorisés. **IA interdite.**
- Durée : 3 à 4 minutes par binôme (avec un temps de parole équilibré).
- Objectif : montrer que vous savez parler d'un sujet en anglais de façon claire, organisée et vivante.

Annexe B — Consignes données aux élèves (exposé 2 avec IA autorisée et encadrée)



1. Consigne :

Choisissez un plat connu d'un autre pays et présentez-le en anglais.

2. Ce que vous devez dire :

- Donnez le nom du plat.
- Dites d'où il vient.
- Décrivez les ingrédients, comment on le prépare et la façon dont on le mange.
- Indiquez où et quand les gens le mangent généralement (à la maison, dans la rue, dans le cadre de fêtes, etc.).
- Expliquez pourquoi vous l'aimez (ou pas).
- Comparez-le avec ce qu'on mange plus généralement chez nous.

3. Étapes à suivre :

- 1) Choisissez un plat célèbre d'un autre pays.
- 2) Rédigez le texte de l'exposé selon un plan simple : introduction – présentation du plat (👉 voir partie 2 : Ce que vous devez dire 👉) – conclusion.
- 3) Présentez le plat à l'oral. Les notes sont autorisées, mais pas un texte entièrement rédigé.

4. Règles importantes :

- Recherches internet et dictionnaire : autorisés.
- IA : autorisée. Tous les prompts utilisés devront être transmis à l'enseignant.
- Durée : 2 à 3 minutes par binôme. Chaque élève doit parler au moins une minute.
- Objectif : montrer que vous savez parler d'un sujet en anglais de façon claire, organisée et vivante.

Annexe C — Mini-guide IA

Mini tuto IA

1) Accéder à l'outil (sans compte) :

1. Ouvre un navigateur Chrome.
2. Va sur : **chatgpt.com**
3. Utilise ChatGPT **sans créer de compte** (mode invité).

2) Ne tape **jamais** dans l'IA des données personnelles ou sensibles, telles que :

- ton nom / prénom, ton adresse, numéro de téléphone, email, etc.
- le nom ou informations de ton école, de ta classe, de ton enseignant, etc.
- des infos sur d'autres élèves (noms, photos, etc.)
- des identifiants, mots de passe, codes, etc.

3) Règles pour ce travail :

- Tu peux utiliser l'IA, mais tu dois **réécrire ton propre script** (≈ une page).
- Tu dois être capable d'expliquer ce que tu présentes.
- Tu dois **transmettre tous les prompts / messages** utilisés.

Partager ta conversation IA :

1. Ouvre la conversation complète.
2. Menu navigateur : **Fichier → Imprimer** (ou **Ctrl+P / Cmd+P**).
3. Destination : **Enregistrer en PDF**.
4. Clique sur Plus de paramètres → puis, dans Format du papier, choisis **Sans titre**.
5. Nom du fichier : **Conversation IA – Prénom 1 Prénom 2**.
6. Enregistre sur le bureau et lève la main.

Annexe D — Questionnaire élèves (version sans IA, Q1–Q11)

H\UTE
ÉC-LE
PÉDAGOGIQUE
BEJUNE

Questionnaire de retour – exposé (1 sur 2)

Prénom : _____

Consigne : pour chaque phrase, indique dans quelle mesure tu es d'accord
(1 = pas du tout d'accord / 10 = tout à fait d'accord).

1. J'ai donné le meilleur de moi-même dans ce travail.

--	--	--	--	--	--	--	--	--	--

1 110

2. J'ai pris du plaisir à faire ce travail.

--	--	--	--	--	--	--	--	--	--

3. J'ai réellement compris le sujet sur lequel portait mon exposé.

--	--	--	--	--	--	--	--	--	--

4. La phase de recherche m'a aidé à construire mes propres idées sur le sujet.

--	--	--	--	--	--	--	--	--	--

5. Je me sentais à l'aise avec les mots et les idées que j'avais préparés avant de présenter.

--	--	--	--	--	--	--	--	--	--

6. Pendant la présentation, mes notes m'ont aidé sans m'empêcher de parler de manière fluide.

--	--	--	--	--	--	--	--	--	--

7. Pendant l'exposé, je savais à tout moment ce que je disais et pourquoi je le disais.

--	--	--	--	--	--	--	--	--	--

8. J'ai eu le sentiment que le discours présenté était vraiment le mien.

--	--	--	--	--	--	--	--	--	--

9. J'ai appris de nouvelles choses sur le sujet que j'ai présenté.

--	--	--	--	--	--	--	--	--	--

10. Je trouve que cette activité évalue des compétences importantes pour mieux parler anglais.

--	--	--	--	--	--	--	--	--	--

11. Cet exposé m'a permis de progresser dans ma manière de parler anglais devant les autres.

--	--	--	--	--	--	--	--	--	--

Annexe E — Questionnaire élèves (version avec IA, Q1–Q21)

H\UTE
ÉC-LE
PÉDAGOGIQUE
BEJUNE

Questionnaire de retour – exposé (2 sur 2)

Prénom : _____

Consigne : pour chaque phrase, indique dans quelle mesure tu es d'accord
(1 = pas du tout d'accord / 10 = tout à fait d'accord).

1. J'ai donné le meilleur de moi-même dans ce travail.

--	--	--	--	--	--	--	--	--	--

1 —————> 10

2. J'ai pris du plaisir à faire ce travail.

--	--	--	--	--	--	--	--	--	--

3. J'ai réellement compris le sujet sur lequel portait mon exposé.

--	--	--	--	--	--	--	--	--	--

4. La phase de recherche m'a aidé à construire mes propres idées sur le sujet.

--	--	--	--	--	--	--	--	--	--

5. Je me sentais à l'aise avec les mots et les idées que j'avais préparés avant de présenter.

--	--	--	--	--	--	--	--	--	--

6. Pendant la présentation, mes notes m'ont aidé sans m'empêcher de parler de manière fluide.

--	--	--	--	--	--	--	--	--	--

7. Pendant l'exposé, je savais à tout moment ce que je disais et pourquoi je le disais.

--	--	--	--	--	--	--	--	--	--

8. J'ai eu le sentiment que le discours présenté était vraiment le mien.

--	--	--	--	--	--	--	--	--	--

9. J'ai appris de nouvelles choses sur le sujet que j'ai présenté.

--	--	--	--	--	--	--	--	--	--

10. Je trouve que cette activité évalue des compétences importantes pour mieux parler anglais.

--	--	--	--	--	--	--	--	--	--

11. Cet exposé m'a permis de progresser dans ma manière de parler anglais devant les autres.

--	--	--	--	--	--	--	--	--	--

12. Avant ce travail, j'utilisais déjà régulièrement ChatGPT (ou une IA similaire).

--	--	--	--	--	--	--	--	--	--

1
▶
 10

13. Avec l'IA, c'était plus facile de m'organiser que sans IA.

--	--	--	--	--	--	--	--	--	--

14. Avec l'IA, j'ai dû moins réfléchir que sans IA.

--	--	--	--	--	--	--	--	--	--

15. Pendant ce travail, l'IA m'a surtout aidé à (choisir 1 ou 2 réponses) :

- ☐ trouver des idées ☐ organiser / structurer ☐ rédiger des paragraphes
☐ corriger l'anglais ☐ simplifier / reformuler / résumer
☐ comprendre le sujet / vérifier ma compréhension ☐ m'entraîner à présenter
☐ autre(s) : _____

16. Sans l'IA, mon exposé aurait été presque identique (1 = complètement différent).

--	--	--	--	--	--	--	--	--	--

17. J'ai modifié moi-même une grande partie du texte / contenu proposé par l'IA.

--	--	--	--	--	--	--	--	--	--

18. J'ai parfois remis en question ce que me disait l'IA. (1 = pas du tout / 10 = très souvent)

--	--	--	--	--	--	--	--	--	--

19. J'ai vérifié certaines informations proposées par l'IA.

--	--	--	--	--	--	--	--	--	--

20. C'était stressant de travailler avec l'IA.

--	--	--	--	--	--	--	--	--	--

21. Je suis resté-e maître de mon travail : c'est moi qui ai décidé ce que je voulais dire.

--	--	--	--	--	--	--	--	--	--

Annexe F — Grille d'évaluation standard (exposé 1)

Grille d'évaluation pour les présentations orales sans IA.

Critères	Mode	Sous-critères	Niveaux de performance	Pts max.	Pts obt.
Contenu présenté		Sujet respecté + informations pertinentes + exemples	Contenu clair, pertinent, avec exemples → 6 ; contenu globalement clair mais partiel → 3-5 ; contenu confus ou hors sujet → 0-2	6	
Organisation		Introduction + ordre logique + conclusion	Organisation claire (intro – idées – conclusion) → 6 ; organisation partielle → 3-5 ; organisation peu claire → 0-2	6	
Langue		Grammaire + vocabulaire + connecteurs	Langue correcte et variée, connecteurs utilisés → 5 ; langue globalement correcte mais limitée → 2-4 ; langue faible ou incorrecte → 0-1	5	
Prononciation & fluidité		Prononciation + débit	Se fait comprendre avec un débit fluide → 4 ; hésitations marquées → 2-3 ; difficile à comprendre ou lecture insistante → 0-1	4	
Attitude de communication		Regard + posture	Communication adaptée à une présentation → 2 ; partiellement adaptée → 1 ; inadaptée → 0	2	
Respect des consignes		Temps + répartition de la parole	Temps respectés → 2 ; léger dépassement ou déséquilibre → 1 ; temps non respectés → 0	2	
			Total :	25	

Annexe G — Grille d'évaluation « IA-informée » (exposé 2)

Grille d'évaluation pour les présentations orales avec IA.

Critères	Mode	Sous-critères	Niveaux de performance	Pts max.	Pts obt.
Contenu présenté		Sujet respecté + informations pertinentes	Contenu clair et pertinent → 4 ; contenu globalement adapté mais incomplet → 2-3 ; contenu confus ou hors sujet → 0-1	4	
Organisation		Introduction + ordre logique + conclusion	Organisation claire (intro – idées – conclusion) → 3 ; organisation partielle → 2 ; peu claire → 1 ; absente → 0.	3	
Langue		Grammaire + vocabulaire + connecteurs	Langue correcte et variée, connecteurs utilisés → 5 ; langue globalement correcte mais limitée → 2-4 ; langue faible ou incorrecte → 0-1	5	
Prononciation & fluidité		Prononciation + débit	Se fait comprendre avec un débit fluide → 3 ; hésitations marquées → 2 ; difficile à comprendre ou lecture insistante → 0-1	3	
Appropriation & compréhension		Réponse aux questions + exemples personnels	Réponses pertinentes + exemple personnel → 5 ; réponses partielles → 2-4 ; réponses vagues ou absence d'exemple → 0-1	5	
Attitude de communication		Regard + posture	Communication adaptée à une présentation → 2 ; partiellement adaptée → 1 ; inadaptée → 0	2	
Respect des consignes		Temps + répartition de la parole + instructions IA	Cadre et temps respectés → 3 ; partiellement respectés → 1-2 ; non respectés → 0	3	
			Total :	25	