



Enseigner le FLE à l'ère de ChatGPT

Usages et mésusage des IA génératives

Serge Bibauw

Université catholique de Louvain

Stage international d'été en didactique du FLE

Louvain-la-Neuve, 14 aout 2024





Le FLE à l'ère de ChatGPT

IA et modèles de langage

Pour l'apprenant

Pour l'enseignant



Le FLE à l'ère de ChatGPT

IA et modèles de langage

Pour l'apprenant

Pour l'enseignant

L'IA n'est pas nouvelle...

- **Intelligence artificielle** = système réalisant des tâches "humaines" : langage, vision, conduite, analyse...
- Tout système qui traite la langue (TAL) ∈ "IA", peu importe si intelligent ou pas : correcteur Word, Google Translate, Siri, ChatGPT...
- "Test de Turing" (1950) amalgame langage et intelligence.
- **ELIZA** : système simulant un psychanalyste (Weizenbaum, 1966)



Welcome to

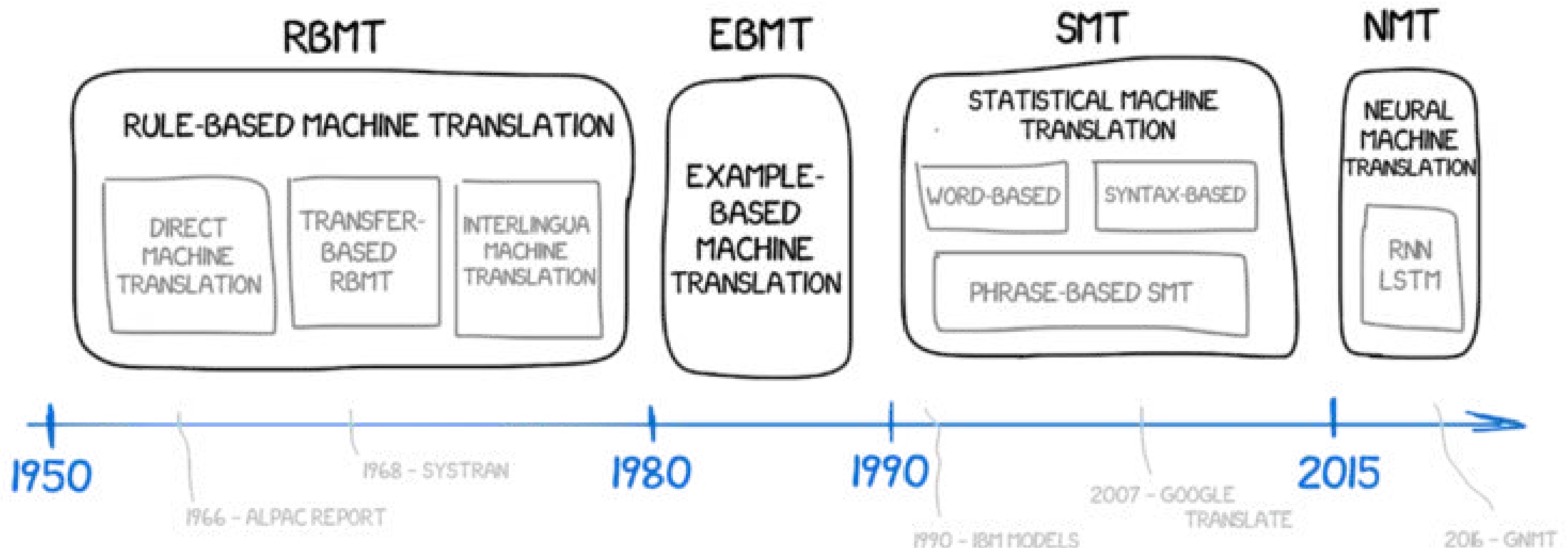
```
EEEEEE LL      IIII  ZZZZZZ  AAAAA
EE      LL      II    ZZ     AA  AA
EEEEEE LL      II    ZZZ    AAAAAAA
EE      LL      II    ZZ     AA  AA
EEEEEE LLLLLL  IIII  ZZZZZZ  AA  AA
```

Eliza is a mock Rogerian psychotherapist.
The original program was described by Joseph Weizenbaum in 1966.
This implementation by Norbert Landsteiner 2005.

```
IZA: Is something troubling you ?
J:   Men are all alike.
IZA: What is the connection, do you suppose ?
J:   They're always bugging us about something or other.
IZA: Can you think of a specific example ?
J:   Well, my boyfriend made me come here.
IZA: Is it important to you that your boyfriend made you come here ?
J:   He says I'm depressed much of the time.
IZA: I am sorry to hear that you are depressed.
J:   It's true. I am unhappy.
IZA: Can you explain what made you unhappy ?
J:
```

...mais les choses ont bien changé !

- Évolutions gigantesques en synthèse vocale, reconnaissance vocale, traduction automatique... et en génération de texte.



Des IA basées sur des règles...

- Règles écrites par des humains
 - p.ex. “Quel temps il fait à Louvain-la-Neuve”
 - SI “Quel temps*”
⇒ “Je ne vous le fais pas dire.”



Welcome to

```
EEEEEE LL      IIII  ZZZZZZ  AAAAA
EE      LL      II    ZZ     AA  AA
EEEEEE LL      II    ZZZ    AAAAAA
EE      LL      II    ZZ     AA  AA
EEEEEE LLLLLL  IIII  ZZZZZZ  AA  AA
```

Eliza is a mock Rogerian psychotherapist.
The original program was described by Joseph Weizenbaum in 1966.
This implementation by Norbert Landsteiner 2005.

```
IZA: Is something troubling you ?
J:   Men are all alike.
IZA: What is the connection, do you suppose ?
J:   They're always bugging us about something or other.
IZA: Can you think of a specific example ?
J:   Well, my boyfriend made me come here.
IZA: Is it important to you that your boyfriend made you come here ?
J:   He says I'm depressed much of the time.
IZA: I am sorry to hear that you are depressed.
J:   It's true. I am unhappy.
IZA: Can you explain what made you unhappy ?
J:
```

Des IA spécifiques à une tâche...

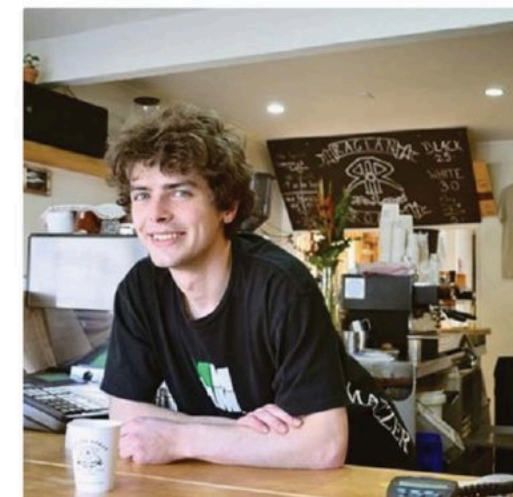
- Ex. : traduction automatique, reconnaissance vocale...
- IA conversationnelle : classification d'intentions (*natural language understanding*)
 - p.ex. "Quel temps il fait à Louvain-la-Neuve ?"
 - → 92%: intent = ask-weather,
21%: intent = ask-location...
- SPELL (Morton, Gunson, & Jack, 2012)
- HALEF (ETS; Sydorenko et al., 2018)
- ...



Good morning! 🌞☁️ On your way to work 🚗 would you mind picking me up some breakfast? Maybe a drink and something to eat. 🍷🍷 Only if it's not too much trouble!

You rush to the coffee shop. Ready to order?

SPEAK WITH THE BARISTA

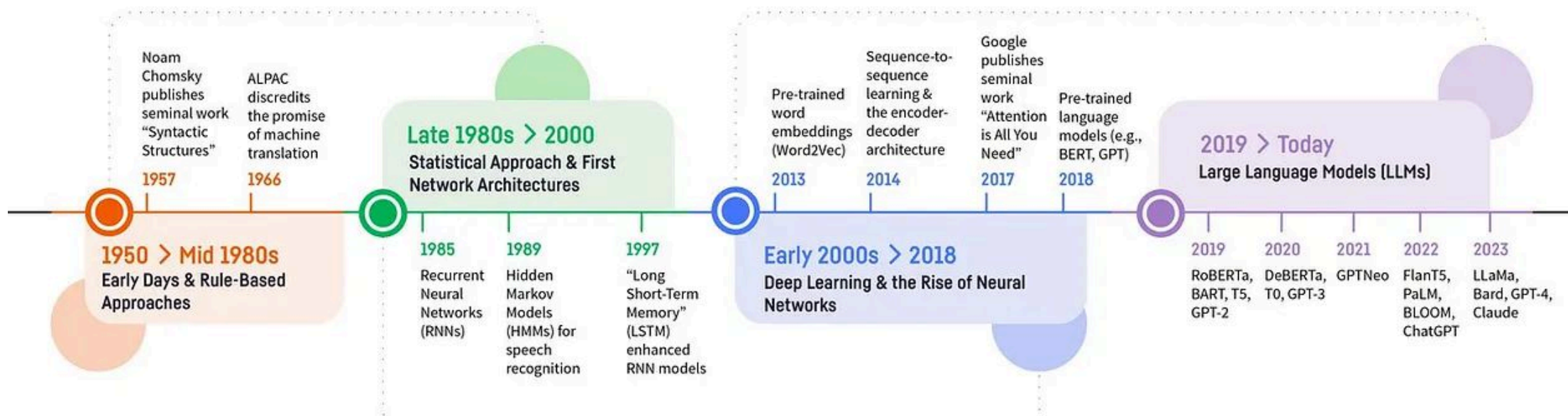


MENU

- Coffee
- Cappuccino
- Latte
- Mocha
- Tea
- Bagel

...aux IA basées sur de grands modèles de langue

- BERT, ChatGPT, GPT-4, Claude, Gemini...
- Deux révolutions :
 - *Transformers* : mécanismes d'attention et fenêtres contextuelles (Vaswani et al., 2017 → BERT)
 - Apprentissage par renforcement + feedback humain (RLHF) (OpenAI, 2022 → ChatGPT)
- Large spectre de tâches possibles.






Grands modèles de langage (LLM)

- **ChatGPT** : modèle de langage génératif.
 - Génère du texte à partir d'un prompt.
 - Pas de mémoire, pas de conscience, pas de compréhension.
 - Pas de modèle du *monde*.
- **Autocomplétion** "sous stéroïdes".
 - Cherche à prédire la suite la plus probable, la plus attendue, pas la plus correcte.
 - Ex. : multiplications à 3 chiffres.
 - Essayons : Mistral AI <https://chat.mistral.ai/> p.ex. $325 * 246$
- **Hallucinations** ≠ failles occasionnelles. Fait partie de leur design !
- LLM = **machine à rêver** dont beaucoup de rêves ressemblent à la réalité.

LLM et hallucinations

- ⇒ **Pas une source de connaissance fiable** pour non-experts.
- Essayez :
 - Quand utilise-t-on “en” ou “à” avec des pays en français ? 🤔
 - Quand s'utilise le passé simple en français ? 🤔
 - Et le présent parfait ? 🤔
- ⚠️ Tendance à multiplier les explications, à en rajouter au-delà du pertinent.
- ⚠️ Segmentation par langue très limitée et dominance de l'anglais dans l'entraînement.
- ⚠️ Pas d'indication d'incertitude ou d'hésitation.
- ⚠️ Admet des aberrations si l'utilisateur insiste.

Mésusages des IA génératives

- Modèle de langage, pas un modèle de monde/vérité.
-  Ne pas considérer une IA générative comme une source de connaissance fiable.
-  Limitez les questions factuelles à celles que l'on peut vérifier.
-  Rester très critique : très convaincant à première vue.



Le FLE à l'ère de ChatGPT

IA et modèles de langage

Pour l'apprenant

Pour l'enseignant

👋 Parlez-vous de ChatGPT à vos élèves ?

1. 🙅 Non, je n'en parle jamais.
2. ✌ Je (ou l'école) l'interdis explicitement à mes élèves.
3. ... Je le recommande à mes élèves.
4. ... J'enseigne à mes élèves comment l'utiliser intelligemment.
5. 👋 Autre.



N'évitez pas le sujet !

- Risque élevé si l'IA reste dans le non-dit.
- Impression pour l'élève que tout le non-dit est interdit.
- Pourquoi triche-t-on ? Pas parce qu'on ne sait pas comment.
- ⇒ En parler, discuter des usages acceptables, expliciter les consignes.
- ⇒ Former à un usage réflexif, expliciter.

AI Assessment Scale (AIAS)

(Perkins et al. 2023)

in *Journal of University Teaching and Learning Practice*, 20 (2)

<https://doi.org/10.53761/qt3azde36>

1	NO AI	<p>The assessment is completed entirely without AI assistance. This level ensures that students rely solely on their knowledge, understanding, and skills.</p> <p>AI must not be used at any point during the assessment.</p>
2	AI-ASSISTED IDEA GENERATION AND STRUCTURING	<p>AI can be used in the assessment for brainstorming, creating structures, and generating ideas for improving work.</p> <p>No AI content is allowed in the final submission.</p>
3	AI-ASSISTED EDITING	<p>AI can be used to make improvements to the clarity or quality of student created work to improve the final output, but no new content can be created using AI.</p> <p>AI can be used, but your original work with no AI content must be provided in an appendix.</p>
4	AI TASK COMPLETION, HUMAN EVALUATION	<p>AI is used to complete certain elements of the task, with students providing discussion or commentary on the AI-generated content. This level requires critical engagement with AI generated content and evaluating its output.</p> <p>You will use AI to complete specified tasks in your assessment. Any AI created content must be cited.</p>
5	FULL AI	<p>AI should be used as a 'co-pilot' in order to meet the requirements of the assessment, allowing for a collaborative approach with AI and enhancing creativity.</p> <p>You may use AI throughout your assessment to support your own work and do not have to specify which content is AI generated.</p>

Déconstruire le rôle « intelligent »

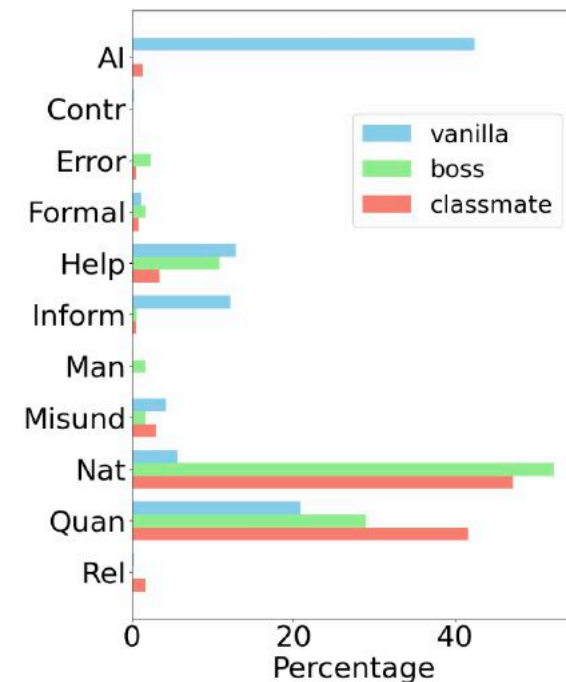
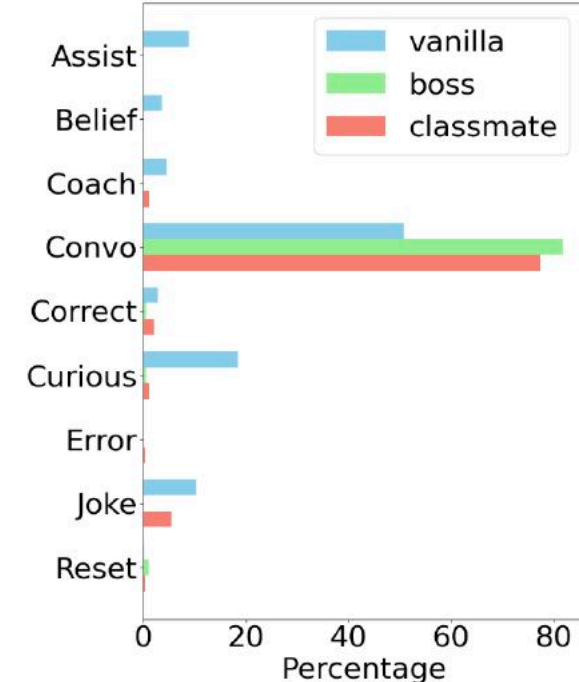
- Si l'IA est « intelligente », le premier rôle de l'apprenant est de la mettre à l'épreuve.
 - Interaction très peu intéressante pédagogiquement au-delà de la curiosité initiale.



ChatGPT Role-play Dataset (Tao, Agrawal, Dombi & Sydorenko, 2024)

- Corpus de dialogues avec apprenants d'anglais
- Quelles postures/demandes des apprenants vis-à-vis de ChatGPT ?
 - Différence si consigne de "tester" ou "jeu de rôle"
 - Si test : peu naturel, peu conversationnel, peu pertinent
 - Si jeu de rôle : plus naturel, plus conversationnel

Analysis	vanilla	boss	classmate
A1: Average conversation length (number of turns)	29.59	14.57	17.11
A2: Average utterance length (Human)	12.18	20.58	19.06
A2: Average utterance length (ChatGPT)	77.66	35.78	46.10
A3: Correlation between human and ChatGPT utterance lengths	0.20	0.14	0.25
A4: Questions as percentage of conversation (Human)	26.34	21.32	21.29
A4: Questions as percentage of conversation (ChatGPT)	14.69	20.34	32.57
A5: Correlation between human questions and number of turns	0.87	0.68	0.51
A5: Correlation between ChatGPT questions and number of turns	0.65	0.77	0.83



Outil pour, mais aussi à apprendre

- Compétence d'interaction avec l'IA sera bientôt aussi essentielle au professionnel que la capacité d'écrire un courriel ou de consulter le web.
 - Compétence fondamentalement langagière
 - Compétence que l'enseignant·e de FLE doit viser
- Requiert que l'enseignant·e lui-même y soit familiarisé·e
 - ⇒ Devenir soi-même un utilisateur critique et réflexif d'Inable'
 - Construction de planifications, séquences, tâches
 - Construction d'exercices et de tests
 - Génération de modèles : textes, dialogues...

Interagir avec l'IA = nouvel objectif communicatif

- “ L'IA ne remplacera pas les humains. Les humains qui utilisent l'IA remplaceront ceux qui ne l'utilisent pas. – Kareem Lakhani ”
 - Interdire: handicap
- Écrire des prompts, etc. = Futurs nouveaux usages (comme le téléphone, l'email, les réseaux sociaux, etc. l'ont été en leur temps).
- Nouvelle responsabilité de l'éducation, particulièrement en langues.

L'IA pour l'apprenant : un soutien à la production (1/2)

- Apprenant comme **producteur** (écrit), IA en soutien
 1. **Génération de modèles** (en amont)
 - Modèles prototypiques idéaux
 - Données à l'apprenant à déconstruire et à réutiliser
 2. **Évaluateur critique et correcteur** des productions de l'apprenant (en aval)
 - Incomplet, mais à visée formative
 - **Feedback** immédiat
 - Problème : par défaut, non progressif
 - Passer par des surcouches (applications ou *prompts* spécifiques)

RECIPE (Han et al, 2023, 2024)

- Interface par-dessus ChatGPT

- Prompt caché : *Act as an English writing class teacher and instruct a student to revise an essay based on the content learned during class. [...] Please guide the student step by step and do not revise or rewrite the essay at once.*

Edit Your Essay Here

Click to check your original essay

A PPL is inevitable in new media environment because it benefits the product company, the program itself, and the viewers. Product company can advertise their products effectively with PPL. In a K-drama "My love from the star", an actress was wearing a jacket and she looked beautiful on it. The jacket instantly went out of stock though it had high price. PPL can enhance the quality of programs. With PPL funding, the program creators can introduce high-quality computer graphics, and hallyu stars in the program. The program will have competitiveness. Naturally, such high quality will lead to fame. Viewers can get authentic information about products through PPL. Since the promoted product is often exposed on screen, viewers can know how the product is used in real life. PPL helps consumers with their decision making process whether to buy the product. PPL is inevitable in new media environment because it can advertise products effectively, enhance the quality of programs and help viewers get authentic information about products. When PPL is properly used, all stake-holders can take advantage of it.

A PPL is inevitable in new media environment because it benefits the product company, the program itself, and the viewers. **First**, product company can advertise their products effectively with PPL. **For example**, in a K-drama "My love from the star", an actress was wearing a jacket and she looked beautiful on it. Thus, the jacket instantly went out of stock though it had high price. **Second**, PPL can enhance the quality of programs. With PPL funding, the program creators can introduce high-quality computer graphics, and hallyu stars in the program. **Therefore**, the program will have competitiveness. Naturally, such high quality will lead to fame. **Third**, viewers can get authentic information about products through PPL. Since the promoted product is often exposed on screen, viewers can know how the product is used in real life. **Thus**, PPL helps consumers with their decision making process whether to buy the product. **To sum up**, PPL is inevitable in new media environment because it can advertise products effectively, enhance the quality of programs and help viewers get authentic information about products. When PPL is properly used, all stake-holders can take advantage of it.

Improve Your Writing with ChatGPT

Hello! Welcome to the main exercise. Can you please tell me what you learned in class so that I can help you better? An example can be:

Today I learned about the comma rules. When there is a list that contains two or more elements, we should use commas to separate them.

Hello! Today, I focused on the proper usage of transition signals. These words help to connect ideas and maintain the flow of the essay.

Great! Let's start revising your essay incorporating the transition signals you learned in class. Can you provide me with the essay you want to revise? We will go through it step by step.

How helpful was the last answer of ChatGPT? *

☹️ 😐 😊 😄 😁

Press Send Button to Ask a Question

L'IA pour l'apprenant : un soutien à la production (2/2)

- Apprenant comme **producteur** (écrit *ou oral*), IA en soutien :

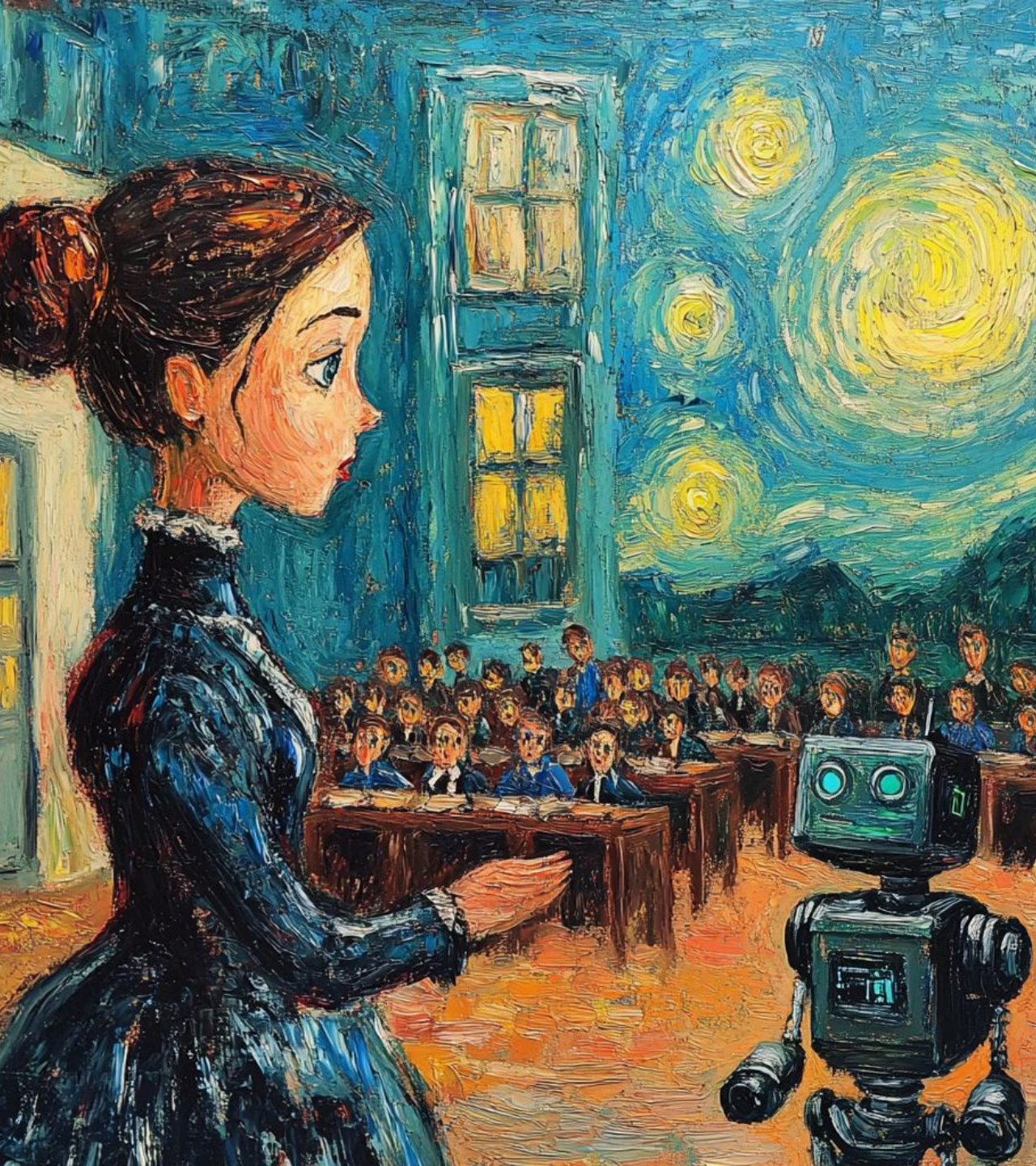
3. Interlocuteur conversationnel

○ Opportunités :

- Pratique centrée sur le sens, particulièrement hors du contexte-cible
- Pratique spontanée écrite ou orale (p.ex., ChatGPT mobile, Talk-to-ChatGPT ou Gemini)
- Moins anxiogène (Alrajhi, 2024; Fathi et al., 2024)
- Potentiellement adaptatif (difficulté), mais à développer

○ Défi :

- Intérêt/motivation/engagement à long terme ? ⇒ Tâches, scénarios pédagogiques
- Pas assez de ruptures conversationnelles (accomodant), pas de feedback par défaut.



Le FLE à l'ère de ChatGPT

IA et modèles de langage


Pour l'apprenant

Sortir du non-dit • Former • Produire

Pour l'enseignant

Planifier • Ressources • Évaluation

Pour l'enseignant

1. Pour planifier
2. Pour créer des textes, des exercices...
3. Pour évaluer ? 

Pour l'enseignant : Planification

- Avant tout pour dépasser l'angoisse de la page blanche.
 - Comme brainstorming.
- Être critique et sélectif !
- Exercice

Pour l'enseignant : Création de ressources

- Générer des textes
 - Générer des audios
 - *ElevenLabs* et autres systèmes de synthèse vocale
 - Générer des vidéos
 - Adapter des textes: raccourcir, transformer, changer le sujet, etc.
- Générer des exercices.
 - QCM, phrases à compléter, ...
 - Ex. *Duolingo English Test* (DET)
- Générer des modèles.

Évaluateur pour l'enseignant ? Prudence !

- Côté enseignant : prudence !
 - Feedback formatif : oui (Pfau, Polio & Xu, 2023)
 - Évaluation sommative : problématique
 - En tout cas pas sans manipulation de la “température” (strict vs. créatif), sinon “régression à la moyenne”
- Côté apprenant : comprendre que, par conception, l'IA ignore/passe par-dessus les incohérences ou problèmes.
 - Ne pas faire traduire pour contrôler la cohérence !

Le FLE à l'ère de ChatGPT

IA et modèles de langage

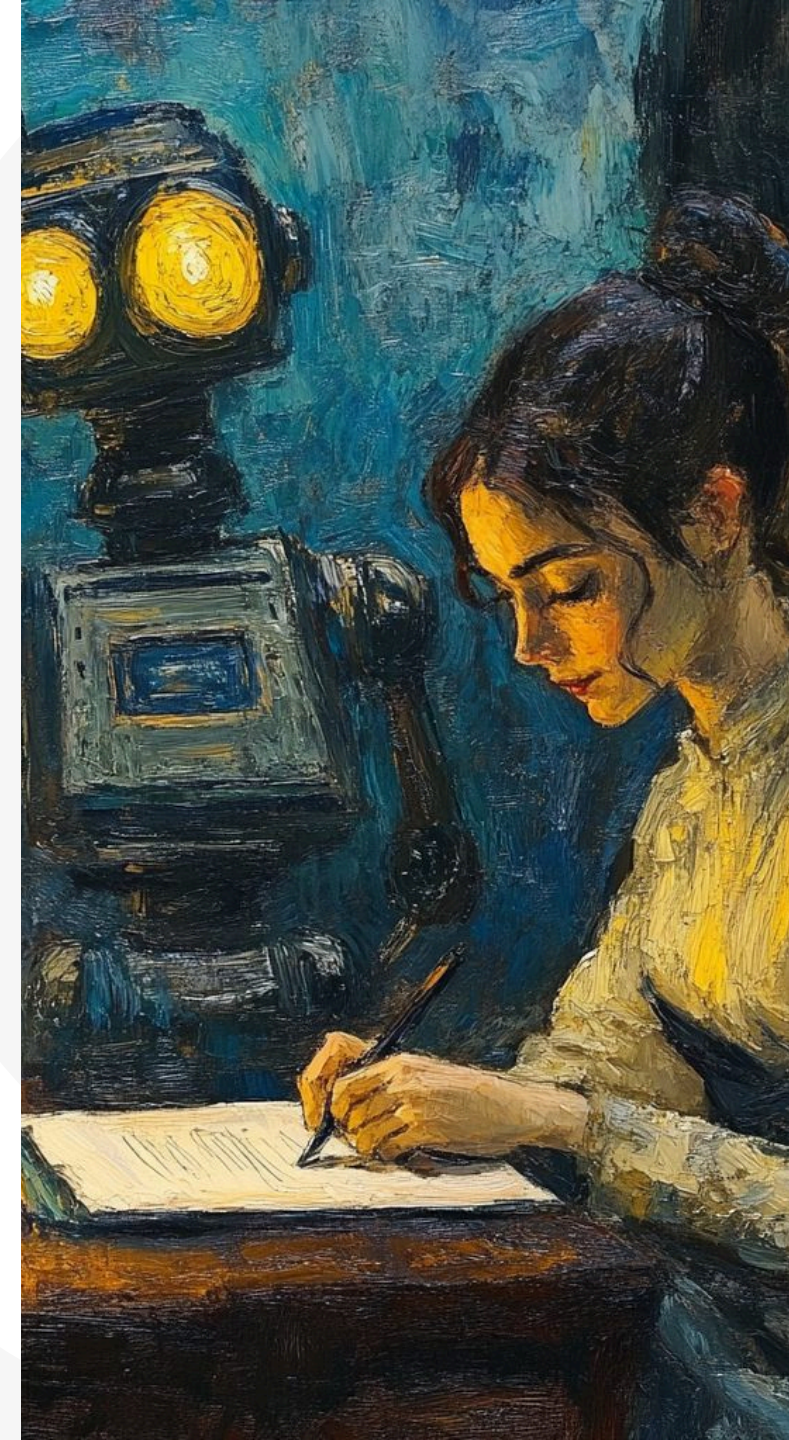
Pas un modèle du monde, hallucinations...

Pour l'apprenant

Sortir du non-dit • Former • Produire

Pour l'enseignant

Planifier • Ressources • Évaluation ⚠



Merci !
Questions ?
Réactions et suggestions bienvenues.

Serge Bibauw
serge.bibauw@uclouvain.be
<https://serge.bibauw.be>



bit.ly/stagefle2024

Illustrations : Midjourney, Unsplash et auteurs respectifs



Sites

Interfaces de chat avec grands modèles de langage

[Mistral AI](#) (IA française)

[Anthropic Claude](#) (capacités similaires à GPT-4)

[Microsoft Copilot](#) (si compte pro, + de respect pour vos données)

Faire tourner un LLM (LLaMA, Mistral...) sur votre ordinateur : [Ollama](#)

Autres IA génératives

Image : DALL-E, Copilot, [Midjourney](#)...

Mise en page : [Copilot Designer](#), [Canva AI](#)

Diagrammes : [Napkin](#)

Voix : [ElevenLabs](#), modèles de synthèse vocale

Vidéo : [Vidnoz](#), [Synthesia](#)...

Utilisation pédagogique

[AI Assesment Scale \(AIAS\)](#).

Version française : [Echelle de l'intelligence artificielle dans l'évaluation](#)

Références

- Alrajhi, A. S. (2024). Artificial intelligence pedagogical chatbots as L2 conversational agents. *Cogent Education*, 11(1), 2327789. <https://doi.org/10.1080/2331186X.2024.2327789>
- Ellis, R., Tanaka, Y., & Yamazaki, A. (1994). Classroom interaction, comprehension, and the acquisition of L2 word meanings. *Language Learning*, 44(3), 449–491. <https://doi.org/10.1111/j.1467-1770.1994.tb01114.x>
- Fathi, J., Rahimi, M., & Derakhshan, A. (2024). Improving EFL learners' speaking skills and willingness to communicate via artificial intelligence-mediated interactions. *System*, 121, 103254. <https://doi.org/10.1016/j.system.2024.103254>
- Han, Z. (2024). ChatGPT in and for second language acquisition: A call for systematic research. *Studies in Second Language Acquisition*, 1–6. <https://doi.org/10.1017/S0272263124000111>
- Jeon, J., Lee, S., & Choe, H. (2023). Beyond ChatGPT: A conceptual framework and systematic review of speech-recognition chatbots for language learning. *Computers & Education*, 206, 104898. <https://doi.org/10.1016/j.compedu.2023.104898>
- Kwon, S. K., Shin, D., & Lee, Y. (2023). The application of chatbot as an L2 writing practice tool. *Language Learning & Technology*, 27(1), 1–19. <https://doi.org/10.1257/73541>
- Morton, H., Gunson, N., & Jack, M. A. (2012). Interactive language learning through speech-enabled virtual scenarios. *Advances in Human-Computer Interaction*, 2012, e389523. <https://doi.org/10.1155/2012/389523>
- OpenAI. (2022, November 30). ChatGPT: optimizing language models for dialogue. *OpenAI Blog*. <https://openai.com/blog/chatgpt>
- Pfau, A., Polio, C., & Xu, Y. (2023). Exploring the potential of ChatGPT in assessing L2 writing accuracy for research purposes. *Research Methods in Applied Linguistics*, 2(3), 100083. <https://doi.org/10.1016/j.rmal.2023.100083>
- Sydorenko, T., Daurio, P., & Thorne, S. L. (2018). Refining pragmatically-appropriate oral communication via computer-simulated conversations. *Computer Assisted Language Learning*, 31(1–2), 157–180. <https://doi.org/10.1080/09588221.2017.1394326>
- Tao, Y., Agrawal, A., Dombi, J., Sydorenko, T., & Lee, J. I. (2024). ChatGPT Role-play Dataset: Analysis of user motives and model naturalness. *LREC-COLING 2024*. <http://arxiv.org/abs/2403.18121>
- Timpe-Laughlin, V., Sydorenko, T., & Daurio, P. (2022). Using spoken dialogue technology for L2 speaking practice: What do teachers think? *Computer Assisted Language Learning*, 35(5–6), 1194–1217. <https://doi.org/10.1080/09588221.2020.1774904>
- Underwood, J. H. (1982). Simulated conversation as a CAI strategy. *Foreign Language Annals*, 15(3), 209–212. <https://doi.org/10.1111/j.1944-9720.1982.tb00248.x>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention Is All You Need (arXiv:1706.03762). *arXiv*. <https://doi.org/10.48550/arXiv.1706.03762>
- Weizenbaum, J. (1966). ELIZA: A computer program for the study of natural language communication between man and machine.