

1. Introduction: Research Experience

Specialised Discourse with Large Language Models? Beyond ChatGPT in Academic English



Josef Schmieid
English Language & Linguistics
Chemnitz University of Technology
https://www.tu-chemnitz.de/phl/english/sections/ling/presentations_js.php
josef.schmieid@phl.tu-chemnitz.de
<http://tinyurl.com/Josef.Schmieid>

1. Introduction: Definition “Discourse”

(OED on line 02/06/24, NOT “specialised”, but d. community, marker, theory, metadiscourse)

Factsheet Meaning & use Etymology Pronunciation Forms Frequency Compounds & derived words

CONTENTS

- The process or faculty of reasoning; reasoned argument or...
- The thread of an argument; a line of reasoning; a reasoned...
- A more or less formal treatment of a subject, in speech or...
- The action or process of communicating thought by means of...
- The quality of being conversant in a subject. Obsolete.
- Interaction, dealings, communication.
- The body of statements, analysis, opinions, etc., relating...
- Linguistics. A connected series of utterances by which...

1951-

8. Linguistics. A connected series of utterances by which meaning is communicated, esp. one forming a unit for analysis; spoken or written communication regarded as consisting of such utterances.

1951 The universe of discourse for each statement in the descriptive analysis is a single whole utterance in the language in question.
Z. S. Harris, *Methods in Structural Linguistics* ii. 11 ...

1995 The matrix language is the language that provides the largest proportion of morphemes within a discourse of at least two sentences.
Lang. in Society vol. 24.476 ...

Show more quotations

📄 Cite 📖 Historical thesaurus ▾
Linguistics

1. Introduction: Research Experience



Funded Research Projects:




- SFB: Identity in Africa: A5 Sociolinguistics of English variation in Africa, International Corpus of English – East Africa (= Kenya, Tanzania)
- Lampeter Corpus of EModE Tracts (1640-1740)
- Internet Grammar of English (incl. stratified English-German Translation Corpus)
- Academic & Journalistic Writing (corpora) (SE Europe, Africa, China)
- SFB/CRC: Hybrid Societies: Humans interacting with embodied digital technologies: D03 Credibility & Language Varieties: Acceptance of conversational pedagogical agents (Chinese, Italian, Czech “Accents”)

1. Introduction: Definition “Discourse”

(<https://www.ldoceonline.com/dictionary/discourse> 02/06/24)

From Longman Dictionary of Contemporary English

dis-course¹ /ˈdɪskɔːs \$-ɔːrɪ/ **noun** *formal*  

- [countable]** a serious speech or piece of writing on a particular subject
discourse on/upon
 a discourse on art
- [uncountable]** serious conversation or discussion between people
 Candidates should engage in serious political discourse.
- [uncountable]** the language used in particular types of speech or writing
 a study of spoken discourse

Examples from the Corpus

discourse

- In many cases, the conventions of academic discourse force researchers to make these assumptions even more explicit and specific.
- the restraints of diplomatic discourse

1. Introduction: Definition "Specialised Discourse"



Specialized Discourse, Linguistic Features and Changing Conventions

Maurizio Gottli (2003)
 Bern: Peter Lang
 Pp. 351
 ISBN 3-03910-027-0 (paper)
 (48.50)

Whenever native and non-native speakers "employ English in a restricted range of social and thematic areas" (Gramley & Pitzold, 1992, p. 246), they use it in order to communicate (specialized) information straightforwardly. The main concern and the major difficulty of linguists who try to describe 'Special Englishes' is that of comparing them to 'General English' and to draw the attention to the differences between the two, which arise at all the levels of language. Lots of work has been done in this direction. Maurizio Gottli's book represents one of the latest attempts to shed light on this topic in a comprehensive way.

1. Introduction: Definition "Discourse"

(<https://dictionary.cambridge.org/dictionary/english/discourse> 02/06/24)

discourse
 noun
 UK /ˈdɪs.kɔːs/ US /ˈdɪs.kɔːrs/

IC or UJ LANGUAGE specialized

NOT academic, metadiscourse!

the use of language to communicate in speech or writing, or an example of this:

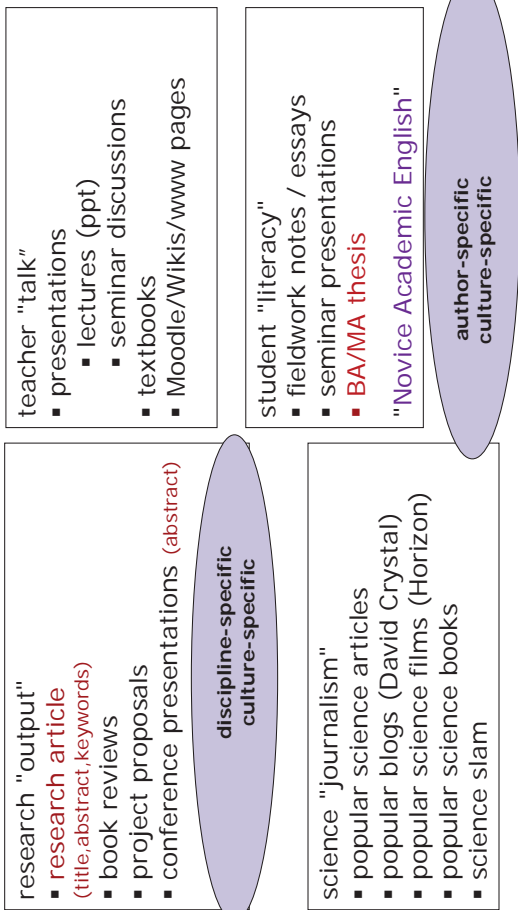
- spoken discourse Ellipsis is a characteristic feature of spoken discourse.
- They may have difficulty in retrieving words and organizing a complex discourse.

See also
 discourse analysis
 discourse marker

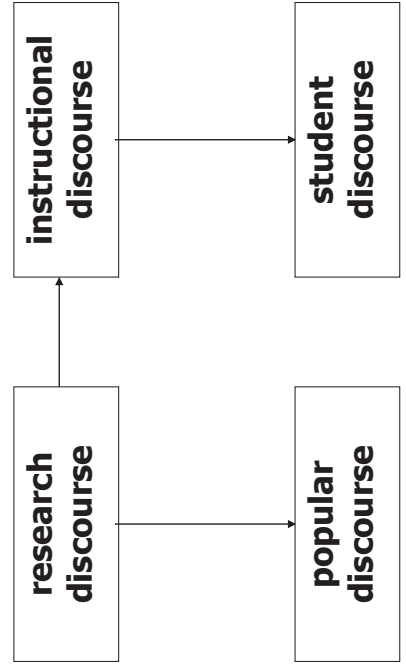
[C or U] formal

- discussion or debate (= formal or political argument), or an example of this:**
- the discourse** On social media, angry and aggressive voices seem to dominate the discourse.
- the public discourse** These ideas have been part of the public discourse for years.
- the political discourse** He accused the media of "dumbing down" the political discourse.
- discourse between / am worried these events could poison discourse between our nations for years to come.**
- discourse with** We need to have a discourse with the public about our work and how it is funded.

1. Introduction: Academic Discourse Genres



1. Introduction: Academic Discourse



1. Introduction: LLM Research Writing Applications

„Artificial Intelligence“

social media

chatbots

Large Language Models (LLMs)

„Office“

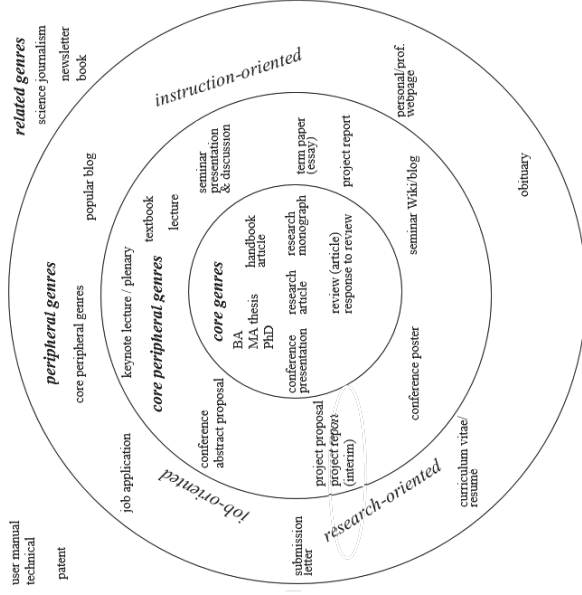
Teacher

Researcher for Ac. Writing (+critical thinking)
NOT Data Processing & Analysing, Outreach, etc.
(cf. Ray 2023, Farrokhnia et al. 2024)

Lecture Planning Course-Curriculum-
Skills Practice „Rules“, etc. Translation

Specialised Concepts Publishing Conventions Style Academic Style
abstract title key words metadiscourse

1. Introduction: Academic Writing Genres



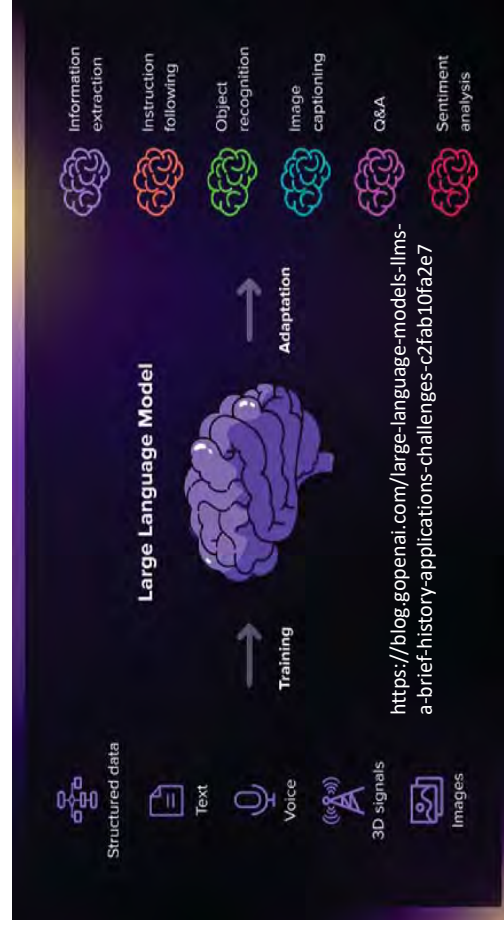
2. Large Language Models: Principles

What Are Large Language Models?

An LLM is a machine learning model that processes and generates human language from training on massive volumes of data. It uses a deep learning algorithm to learn the patterns and nuances of language and provide human-like responses to a wide variety of queries and prompts. LLMs can also learn new skills such as creative writing or sentiment analysis.

The “large” in its name refers to the size of the datasets on which it’s trained. The goal of an LLM at its core is to infer what comes next in a sequence of texts. ChatGPT is a conversational agent built on top of an LLM that can generate text in a way that’s extremely similar to humans. When it comes to expanding the search experience, Coveo is testing **generative AI** capabilities that surface answers directly within search results.

2. Large Language Models: Principles



<https://blog.gopenai.com/large-language-models-llms-a-brief-history-applications-challenges-c2fab10fa2e7>

2. Large Language Models: P=Reward Model Training

Use of pre-trained neural net Transformer models on human language management

Generative Pretrained Model

ChatGPT V1

Labelled data

1 Fine-tune by training on labelled data in supervised mode

2 Generating data with supervised mode using expert instances of ChatGPT V1

3 Humans will rank these answers from bad to very good

4 A Reward Model trained with these responses ranked by human experts, will be used for ChatGPT to score its answers during the Reinforcement Learning training.

Reinforcement Learning using the already trained Reward Model

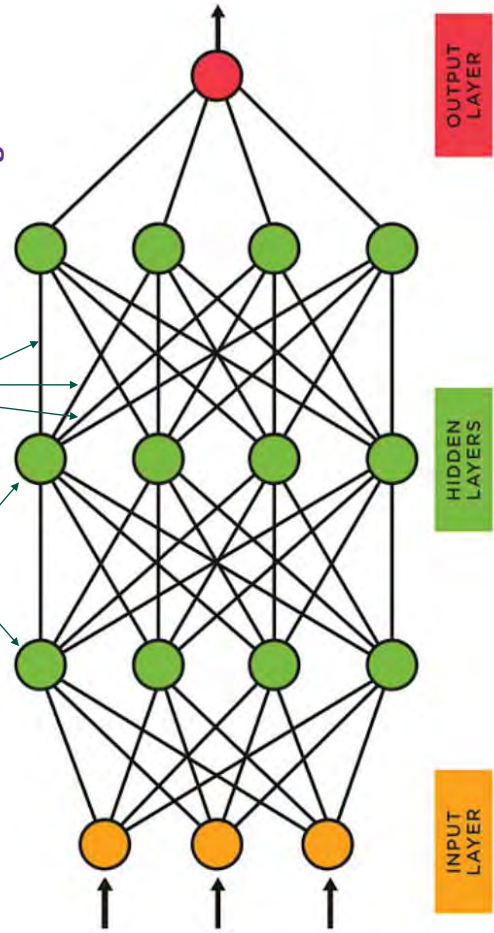
ChatGPT Final

The training of LLMs has become controversial when it is outsourced to "clickworkers" with low wages (e.g. to detect "toxic language"):
<https://time.com/6247678/openai-chatgpt-kenya-workers/>
<https://www.hopeforthefuture.at/en/clickworkers-exploited-for-artificial-intelligence/>
 Generally, all users are asked to give Feedback on (alternative) LLM answers (ChatGPT opt-in):
 e.g. Gemini asks: Good Answer Bad Answer Adapt Answer

P=pre-trained

2. Large Language Models: G=Neural Network

Cf. human brain = 100 billion neurons * 1000 synapses



G=generative

2. Large Language Models: Linguistic Principles „MI“

English-Corpora.org

corpora PDF guides videos related resources users my account upgrade help

In our corpora, Mutual Information is calculated as follows:
 $MI = \log \left(\frac{AB * sizeCorpus}{(A * B * span)} \right) / \log(2)$

Suppose we are calculating the MI for the collocate *color* near *purple* in BNC.
 A = frequency of node word (e.g. *purple*): 1268
 B = frequency of collocate (e.g. *color*): 312
 AB = frequency of collocate near the node word (e.g. *color* near *purple*): 22
 sizeCorpus = size of corpus (e words; in this case the BNC): 96,263,399
 span = span of words (e.g. 3 to left and 3 to right of node word): 6
 $\log(2)$ is literally the natural log of the number 2: 0.6931

$MI = 11.30 = \log \left(\frac{22 * 96,263,399}{(1246 * 112 * 6)} \right) / 0.6931$

BNCweb

HELP	WORDS	INFO	ALL	US
1	LOOSESTRIFE	7	87.50	13.46
2	BERBERS	7	36.84	12.21
3	TYRAN	5	35.71	12.17
4	SANDPAPERS	5	26.47	11.73
5	COLOR	22	19.64	11.30
6	CRIMSON	13	3.69	8.81
7	MAUTE	7	3.25	8.73
8	BRUDE	6	1.84	8.71
9	HAZE	3	3.26	8.25
10	ROBES	5	2.91	7.79
11	PINK	43	3103	7.48
12	CRUMPER	13	1202	7.28
13	PACK	9	852	7.08
14	PROSE	9	852	7.08
15	FOULAGE	7	706	7.00

2. Large Language Models: T=Transformer Visualisation

LLM Visualization

Chapter: Output

T=transformer

<https://bbycroft.net/llm>

Table of Contents

- Intro
- Introduction
- Preliminaries
- Components
- Embedding
- Layer Norm
- Self Attention
- Projection
- MLP
- Transformer
- Softmax
- Output

LLM Visualization

2. Large Language Models: SWOT Analysis ChatGPT in education



Farrokhnia et al. 2024: 468

3. Brainstorming: Ask ChatGPT, who?! can do anything

ChatGPT

Sure, here are some examples of things you can tell me to do:

1. Generate a poem about nature.
2. Summarize the plot of "To Kill a Mockingbird" by Harper Lee.
3. Provide a list of the top 10 tourist attractions in Paris.
4. Explain the concept of quantum entanglement.
5. Create a simple recipe for chocolate chip cookies.
6. Generate a random 5-digit number.
7. Translate the phrase "Hello, how are you?" into Spanish.
8. Summarize the key points of a TED Talk on artificial intelligence.
9. Describe the process of photosynthesis in plants.
10. Provide a brief biography of Albert Einstein.

Which task would you like to start with?



2. Large Language Models: Linguistic Principles „collocates“

alcohol (NOUN) **COLLOC**

See in Web Collocates Clusters Topics KWIC HELP

TOPICS (more)

drinking, drug, liquor, substance, cocaine, abuse, smoking, addiction, marijuana, heroin, smoking, consumption, drink, beer, drunk, treatment, excessive, use, harmful

COLLOCATES (more)

NOUN drug, use, abuse, consumption, tobacco, problem, blood, level

VERB drink, consume, rub, avoid, abuse, smell, smoke, associate

ADJ fetal, excessive, illegal, heavy, moderate, illicit, smoking, related

ADV e.g. be, prenatally, voluntarily, negatively, excessively, moderately, capita

RELATED WORDS

alcoholic, alcoholicism, alcohol-related, alcohol-free, alcohol-based, blood-alcohol

SEARCH **WORD** **CONTEXT** **ACCOUNT**

BLOG WEB TV/M SPOK PIC MAG NEWS ACAD

1. a liquor or brew containing alcohol as the active agent 2. any of a series of volatile hydroxyl compounds that are made from hydrocarbons by distillation

Youdfish PlayPhrase Yam DE: Google WordNet Reverso Linguee

3. Brainstorming: Awareness of AI since ChatGPT

Mizumoto/Eguchi (2023):

The world was taken by surprise by the release of ChatGPT (<https://chat.openai.com/>) at the end of the year 2022. ChatGPT is a chatbot that answers human questions with an AI that seems to have a perfect understanding of the language. ...

The emergence of an AI tool capable of comprehending and producing a language like ChatGPT marks the beginning of a new era in which humans and AI coexist in L2 language learning and teaching, as well as in its research.

However, the use of ChatGPT is currently being evaluated, and some academic journals have cautioned against its irresponsible application (e.g., [Nature Editorial, 2023](#)). Despite criticisms from scholars like Chomsky, who denounce ChatGPT as a form of "high-tech plagiarism" ([EduKitchen, 2023](#)), it is evident that AI language models are here to stay, and it is our responsibility to educate students on their ethical use because "the future is now" ([Pavlik, 2023](#), p. 9). At the same time, learners, teachers, and researchers can benefit from acquiring effective methods for utilizing GPT in their work, particularly non-native English speakers, who can be linguistically empowered. Language models such as GPT can be best understood as an ever-present, logical assistant, as AI is not currently a substitute for human expertise.

3. Brainstorming: Ask Gemini ethical issues

Large Language Models (LLMs) are powerful tools, but their use in research articles raises some ethical concerns. Here are some key areas to consider:

Authorship and Credit: Who gets credit for research findings when an LLM helps generate content? Should the LLM be a co-author? How much human oversight is needed? These are open questions that can blur the lines of traditional authorship. Oxford study on LLM responsibility [invalid URL removed]

Transparency and Reproducibility: Research relies on being able to repeat and verify findings. If an LLM is involved in data analysis or literature review, how can researchers ensure transparency about the LLM's role and limitations? Can others replicate the research if they don't have access to the same LLM?

Bias and Accuracy: LLMs are trained on massive amounts of data, which can reflect societal biases. These biases can be unknowingly incorporated into research articles, leading to skewed or inaccurate findings. Researchers need to be aware of these potential biases and take steps to mitigate them.

Overtrust and Misinformation: It's easy to be impressed by an LLM's ability to generate text. However, LLMs can still produce factually incorrect or misleading content. Researchers need to critically evaluate the LLM's outputs and avoid uncritically accepting them as truth.

These are just some of the ethical considerations surrounding LLM use in research. As this technology continues to develop, researchers and publishers will need to grapple with these issues and develop best practices for ethical and responsible LLM use.

3. Brainstorming: Ask Claude 3 Sonnet specifically1

As an AI assistant, I can help you with various aspects of writing a research article:

1. **Literature Review**
2. **Structuring the Article:** I can provide guidance on the typical structure of a research article, including the introduction, literature review, methodology, results, discussion, and conclusion sections. I can help you organize your ideas and ensure a logical flow of information.
3. **Writing and Editing:** I can help you with the actual writing process, from drafting paragraphs and sections to improving clarity, coherence, and language usage. I can also assist with editing and proofreading to ensure your article is well-written and free of errors.
4. **Formatting and Referencing**
5. **Data Analysis and Visualization**
6. **Manuscript Submission**
7. **Collaboration and Feedback**

It's important to note that while I can provide valuable assistance, I cannot replace your subject-matter expertise or critical thinking. My role is to support you throughout the writing process, but the research, analysis, and conclusions should ultimately come from you and your co-authors.

3. Brainstorming with Bard (=Gemini): writing for journals

write on digital dissemination practices for linguists for Publication in Peer-Reviewed Journals

Vorschlag 1
Sure, here are some digital dissemination practices for linguists for publication in peer-reviewed journals:

informal style

Vorschlag 2
Sure, here are some digital dissemination practices for linguists for publication in peer-reviewed journals:

Vorschlag 3
Sure, here are some digital dissemination practices for linguists for publication in peer-reviewed journals:

choice of options

Sure, here are some digital dissemination practices for linguists for publication in peer-reviewed journals:

- Publish in open access journals: Open access journals make their articles freely available online, which can help to increase the visibility and impact of research. There are a number of open access journals that specialize in linguistics, such as the *Journal of Linguistics* and the *International Journal of Language and Communication Disorders*.
- Create and share research data: Linguists can share their data through online repositories, such as the Linguistic Data Consortium and the Open Linguistics Archive. This can make it more accessible to other researchers and the public, and can facilitate further research on the data.
- Use social media: Social media platforms like Twitter, Facebook, and LinkedIn can be used to share research findings with a wider audience. Linguists can create accounts for themselves and their research groups, and can use these accounts to post links to their articles, data, and other resources. They can also use social media to interact with other linguists and the public, and to answer questions about their research.
- Create and maintain a website: Linguists can create websites to showcase their research and make it easier for people to find their work. The

3. Brainstorming: Ask ChatGPT 4 ethical issues

Using Large Language Models (LLMs) for research articles raises several ethical considerations that researchers should carefully address. Here are some key ethical issues:

1. Bias and Fairness
2. Plagiarism
3. Authorship and Attribution
4. Privacy and Confidentiality
5. Manipulation and Misinformation
6. Impact on Academic Publishing
7. Environmental Impact
8. Equitable Access

Addressing these ethical issues requires a combination of **awareness, transparency, accountability, and proactive measures** to mitigate potential risks and promote responsible use of LLMs in research articles.

4. Research Publishing: GPT4 Research Article Structure

Research articles in digital humanities typically follow a structure that combines traditional humanities scholarship with innovative digital methodologies and tools. While the specific structure may vary depending on the nature of the research and the preferences of individual researchers or journals, here's a prototypical structure commonly used in digital humanities research articles:

1. Title Page
 2. Abstract
 3. Keywords
 4. Introduction
 5. Literature Review
 6. Methodology
 7. Case Study or Analysis
 8. Discussion
 9. Conclusion
 10. References
 11. Appendices (if applicable)
- [https://token.ujk.edu.pl/volumes/volume-16/TOKEN 16 \(2023\)](https://token.ujk.edu.pl/volumes/volume-16/TOKEN 16 (2023))
Special issue: Patterns of Language Variation and Change in Academic Writing
Guest editors: Josef Schmied (*Chemnitz University of Technology*), Marina Bondi (*University of Modena and Reggio Emilia*), Olga Dontcheva Navratilova (*Masaryk University*) and Carmen Pérez-Llantada (*University of Zaragoza*)

This structure provides a framework for organizing research articles in digital humanities, integrating traditional humanities scholarship with innovative digital methodologies and technologies. Researchers should **adapt and refine this structure to fit the specific requirements** and conventions of their discipline, as well as the aims and scope of their research projects.

4. Research Publishing: ChatGPT academic style

Consensus GPT excels in a variety of research applications, including:

- **Real Study Insights:** Ask Consensus to interpret findings from genuine studies on topics like the cognitive benefits of creatine.
- **Enhanced Academic Writing:** Utilize Consensus to draft literature reviews or academic pieces with valid citations.
- **Targeted Paper Searches:** Effortlessly locate specific research papers without relying on precise keyword matches.
- **Advanced Filtering:** Apply filters to refine searches based on publication type, journal prestige, and study design.
- **Research Paper Analysis:** Have Consensus examine and summarize key elements of a research document.
- **Concept Extraction:** Request a compilation of strategies or findings across multiple studies.
- **Science-based Content Creation:** Draft authoritative blogs on health and wellness topics grounded in scientific evidence.

<https://consensus.app/home/blog/introducing-researchgpt-by-consensus/>

We are excited to announce the official launch of Consensus, formerly known as ResearchGPT, as the leading GPT product available in OpenAI's GPT marketplace. We're not just a GPT. Did you know you can try the full Consensus search engine experience for free at <https://consensus.app/home/>?

4. Research Publishing: GPT4o Conclusions (without analysis?)

Based on the findings from the case studies and the literature review, several conclusions can be drawn regarding the utilization of Large Language Models (LLMs) in research article composition:

1. **Enhanced Writing Efficiency and Productivity:**
 - LLMs offer researchers a powerful tool for streamlining the writing process, leading to increased efficiency and productivity.
 - Researchers using LLMs can generate coherent and contextually relevant text more quickly, enabling them to produce research articles in a more timely manner.
2. **Facilitated Interdisciplinary Collaboration**
3. **Improved Writing Creativity and Exploration**
4. **Ethical Considerations and Challenges**
 - The use of LLMs in research article composition raises important ethical considerations related to bias mitigation, plagiarism prevention, and authorship attribution.
 - Ethical guidelines and safeguards are essential to ensure responsible use of LLMs and mitigate potential risks associated with bias, plagiarism, and data privacy.
5. **Opportunities for Future Research and Development**

In conclusion, the integration of LLMs into research article composition holds immense potential for enhancing writing efficiency, promoting interdisciplinary collaboration, and fostering creativity in the scholarly community. However, it is crucial for researchers to navigate the ethical considerations associated with LLM-assisted writing responsibly and transparently. By critically evaluating the opportunities and challenges of LLMs in academic writing, researchers can harness their transformative potential to advance knowledge and scholarship in the digital age.

4. Research Publishing: the Art of Prompting

1. **Title**
Suggest a title for this research journal abstract
2. **Abstract**
Suggest a 300 word abstract for this title: Using large Language Models to write Research articles
3. **Keywords**
Suggest 5 keywords for this research journal abstract
4. **Introduction**
Write an (historical?) introduction for this research journal abstract
5. **Literature Review**
Write a literature review for this research journal abstract
6. **Methodology**
Which methodologies can I use for this article?
7. **Case Study or Analysis**
Find 3 case studies for this issue
8. **Discussion**
9. **Conclusion**
Which conclusions may I arrive at in this article?
10. **References**
Which 10 references are crucial for this article?

4. Research Publishing: Improve AI Proposals

DeepL English (britisch) Wörterbuch Änderungen anzeigen 58

Optimierte Texte mit DeepL Write

The fast development of AI opens up new possibilities for learning and using foreign languages. Although tools like ChatGPT, Bard, Perplexity, or DeepL are currently free, their applications need to be learned to use them effectively for personal or professional needs. (often called "stochastic parrots"). We'll apply these tools to various types of English texts, such as travelogues, book reviews, summaries, etc., and discuss the results generated by AI. We don't write - we improve. We don't translate - we compare versions. This course is suitable for anyone who wants to create foreign-language texts, especially at an advanced level, either for personal or professional purposes.

In this course we will explain the basics of the most common AI tools (often called "stochastic parrots"). We'll apply these tools to different types of English texts, such as travel reports, book reviews, summaries, etc., and discuss the results generated by the AI. We don't write - we improve. We don't translate - we compare versions. This course is suitable for anyone who needs to produce foreign language texts, especially at an advanced level, either for personal or professional purposes.

4. Research Publishing: Personalising ChatGPT advice

how can I express my academic identity in my academic English through metadiscourse?

good concepts
typical examples
no references

Expressing your academic identity in your academic writing, especially through metadiscourse, is crucial for establishing credibility, engaging readers, and conveying your authority on the subject. Metadiscourse refers to the language used to guide readers through your text and convey your attitude and stance toward the material. Here are some strategies for using metadiscourse effectively to express your academic identity in academic English:

- Hedging and Boosting:** Use hedging to indicate uncertainty or caution, and boosting to emphasize your confidence in your claims. For example:
 - Hedging: "It appears that the results may suggest a correlation..."
 - Boosting: "The robust findings unequivocally support the hypothesis..."
- Engagement Markers:** Engage with your readers by addressing them directly or signaling your awareness of their presence. This can foster a sense of connection and collaboration. For instance:
 - "In this paper, we will demonstrate..."
 - "As you can see from the data..."
- Epistemic Modality:** Indicate your degree of certainty or belief in the information you're presenting. Use adverbs and modal verbs like "clearly," "possibly," "likely," or "certainly" to express your stance:
 - "This finding clearly demonstrates..."
 - "It is likely that this phenomenon occurs because..."

4. Research Publishing: Evaluate AI Proposals

English (britisch) Wörterbuch Änderungen anzeigen 58

Vorschläge Varianten

The rapid development of AI is opening up new possibilities for learning and using foreign languages. Although tools like ChatGPT, Bard, Perplexity, or DeepL are currently free, their applications need to be learned to use them effectively for personal or professional needs. (often called "stochastic parrots"). We'll apply these tools to various types of English texts, such as travelogues, book reviews, summaries, etc., and discuss the results generated by AI. We don't write - we improve. We don't translate - we compare versions. This course is suitable for anyone who wants to create foreign-language texts, especially at an advanced level, either for personal or professional purposes.

The rapid development of artificial intelligence is opening up new possibilities for learning and using foreign languages. Although tools like ChatGPT, Bard, Perplexity, or DeepL are currently free, their applications need to be learned to use them effectively for personal or professional needs. (often called "stochastic parrots"). We'll apply these tools to various types of English texts, such as travelogues, book reviews, summaries, etc., and discuss the results generated by AI. We don't write - we improve. We don't translate - we compare versions. This course is suitable for anyone who wants to create foreign-language texts, especially at an advanced level, either for personal or professional purposes.

The rapid development of AI presents new opportunities for learning and utilising foreign languages. While tools such as ChatGPT, Bard, and Perplexity are currently available at no cost, it is necessary to gain an understanding of their applications in order to utilise them effectively for personal or professional purposes.

This course will present fundamental concepts associated with the most prevalent AI tools, which are often referred to as "stochastic parrots". The course will apply these tools to a variety of English texts, including personal reviews, summaries, and others. The results generated by AI will be discussed. Rather than writing, the focus will be on improving existing texts. Instead of translating, the objective will be to compare versions. This course is suitable for anyone who wants to create foreign-language texts, especially at an advanced level, for personal or professional purposes.

Stil akademisch

5. Outlook: Specialised Discourse, Prototypes vs. Author Identity

Prototypical (culture-, discipline-specific) expectations vs. Author Identity

→ adapt

- Set specific academic style: technical terms + metadiscourse
- Customise voice (not Standard US)

Several large language model (LLM) applications allow users to upload their own writings to teach the model their personal style. These applications offer fine-tuning or customization options, which can be especially useful for personalizing outputs. Here are a few prominent ones:

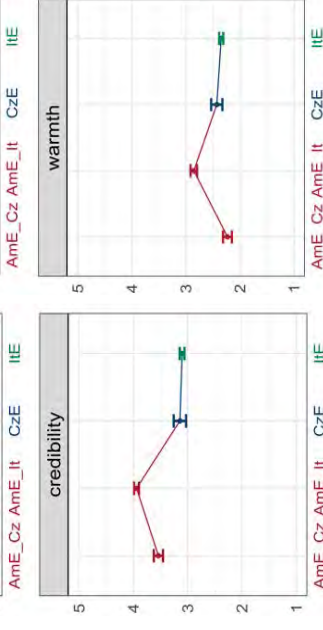
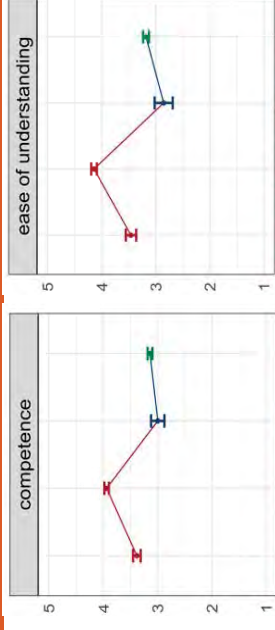
- OpenAI's GPT-3 and GPT-4 (via API and Playground):**
 - OpenAI allows users to fine-tune models using their own data. By providing a set of examples of your writing, you can customize the output to better match your personal style.
 - Playground:** Offers a user-friendly interface for testing and tweaking model outputs with your custom prompts and data.
 - API:** Provides a more extensive setup for fine-tuning and deploying customized models in various applications.

5. Outlook: Reading of Personalised Agents and Credibility

Non-native English lecture



Italian and Czech students give lower rates to their own varieties of English than to American English, for competence and warmth



6. Conclusion: The human role remains essential

ChatGPT thinks human judgment remains essential (Zappavigna 2023)

for

1. Defining the task
2. Data pre-processing
3. Evaluation and refinement
4. Interpretation and analysis

“Generative AI and the end of corpus-assisted data-driven learning? Not so fast!” (Crosthwaite & Baisa 2023)

- Knowing the data
- Authenticity
- Replicability
- Multimodality
- Safety
- Hallucinations
- Active vs. passive learning

“traditional” reliable resources are available online

5. Outlook: Lecture by Agents with Personalised Voice

Research Context: Conversational Pedagogical Agents (CPA) in Hybrid Societies (cf. Ivanova, Albrecht, Schmieid fc.).

Mission

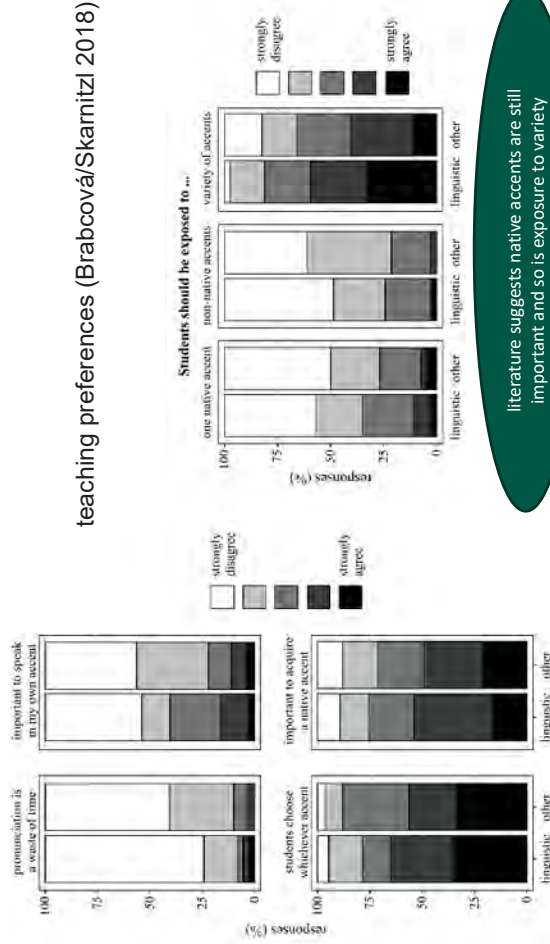
In hybrid societies, humans and embodied digital technologies should interact as seamlessly as humans among each other.

RQ1 Which specific non-native linguistic cues of CPAs influence the learning performance of non-native human learners?

RQ2 Which specific non-native linguistic cues influence attributed credibility and acceptance of CPAs by non-native human learners?

RQ3 How much does a linguistically credible CPA influence the learning performance in non-native educational contexts?

5. Outlook: Reading of Personalised Agents and Credibility



6. Conclusion: AI detection tools

100% AI likeness score

Top sentences driving AI probability

- 1 It helps to soften assertions and make them more diplomatic. **0.41**
- 2 In business texts, hedging is commonly employed to maintain a sense of professionalism and objectivity, and to acknowledge uncertainty or variability in certain situations. **0.38**
- 3 Hedging is a linguistic phenomenon where cautious or non-committal expressions are used to avoid making absolute statements or to mitigate potential risks or uncertainties associated with statements. In business texts, hedging is commonly employed to maintain a sense of professionalism and objectivity, and to acknowledge uncertainty or variability in certain situations. It helps to soften assertions and make them more diplomatic. **0.38**

References

Aktay S., S. Gök, D. Uzunoglu (2023). ChatGPT in education. *Türk Akademik Yayınlar Dergisi (TAY Journal)* 7(2), 378-406.

Brabová, K., R. Skarmitz (2018). Foreign or native-like? The attitudes of Czech EFL learners towards accents of English and their use as pronunciation models. *Studie z aplikované lingvistiky. Studies in Applied Linguistics* 1, 38-50.

Claude AI (2024). How to write a research paper with Claude AI. <https://claude.ai.uk/how-to-write-a-research-paper-with-claude-ai/>

Crosthwaite, P., & Baisa, V. (2023). Generative AI and the end of corpus-assisted data-driven learning? Not so fast! *Applied Corpus Linguistics*, 3(3), 100066.

Farrokhnia, M., Banihashem, S. K., Noroozi, O., & Wals, A. (2024). A SWOT analysis of ChatGPT: implications for educational practice and research. *Innovations in Education and Teaching International* 61(3), 460-474. <https://doi.org/10.1080/14703297.2023.2195846>

Nanova, M., Albrecht, S., & Schmieid, J. (f.c.). Learning outcomes with Text-to-Speech synthesis of native and non-native English varieties. *Johannesburg, P.* (2024). *Writing Your Thesis with ChatGPT: Research, Scholarship and Academic Writing in the Age of Generative AI.* <https://writingyourthesiswithchatgpt.wordpress.com/>

Mizumoto, A., M. Eguchi (2023). Exploring the potential of using an AI language model for automated essay scoring. *Research Methods in Applied Linguistics* 2, 100050

Ray, P. P. (2023). ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet of Things and Cyber-Physical Systems* 3, 121-154, ISSN 2667-3452, <https://doi.org/10.1016/j.iotcps.2023.04.003>.

Skrabut, St. (2023). *80 Ways to Use ChatGPT in the Classroom: Using AI to Enhance Teaching and Learning.* StansKrabut.

Schmieid, J., M. Bondi, O. Dontcheva Navratilova, & C. Pérez-Llantada (eds.) (2023). Patterns of Language Variation and Change in Academic Writing. Special Issue. *TOKEN: A Journal of English Linguistics* 16. <https://token.ujk.edu.pl/volumes/volume-16/>

Schmieid, J. (2023). Confidence and trust in online academic discourse: Integrating new technologies into teaching and learning. Schmieid, J. & M. Ivanova (eds. 2023). *Comparing Confidence and Trust On- and Offline.* Göttingen: Cuvillier, 1-14.

Zappavigna, M. (2023). Hack your corpus analysis: How AI can assist corpus linguists deal with messy social media data. *Applied Corpus Linguistics* 3(4), 100067.

6. Conclusion: Ethical and legal issues

In the end the AI user decides – and is responsible!

- Copyright issues and unclear! <https://noyb.eu/en/chatgpt-provides-false-information-about-people-and-openai-cant-correct-it>
- In academic writing and publishing all aids and resources used must be specified clearly!

KD nuggets

After the launch of ChatGPT, the Pandora box opened. We are now observing a technological shift in the ways we do work. People are creating websites, apps, and even writing novels using ChatGPT.

With all the hype and introduction of AI generative tools, we have seen a rise in bad actors. If you are following the latest news, you must have heard that ChatGPT has passed the Wharton MBA exam. There is a whole list of exams passed by ChatGPT, from medical to a law degree - List Here. Are the Exams ChatGPT Has Passed so Far?

Apart from exams, students are using it to submit assignments, writers are submitting generative content, and researchers are producing high-quality papers just by typing prompts.

To counter the abuse of generative content, I am introducing you to 5 free AI content detector tools. You can use them to check the validity of the content and improve your SEO ranking.

<https://www.kdnuggets.com/2023/02/5-free-tools-detecting-chatgpt-gpt3-gpt2.html>

6. Conclusion: AI detection and un-detection – the race is on

AI Detector Score

READS LIKE AI!

Want Undetectable AI Content? Then AI Humanizer Has You Covered!

The AI Humanizer uses a proprietary mix of 3 AI engines, NLP and semantic analysis algorithms to rewrite sentences and paraphrase paragraphs in a way that is so humanlike it bypasses even the toughest and most accurate AI detection tools - like this one!

Characters: 434 / 2500

Did you write this yourself or you're going to get grade optimized?

Fully Human

Rewrite All AI Sentences