Critical Discourse Studies

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/rcds20

The rise of large language models: challenges for Critical Discourse Studies

Mathew Gillings, Tobias Kohn & Gerlinde Mautner

To cite this article: Mathew Gillings, Tobias Kohn & Gerlinde Mautner (11 Jul 2024): The rise of large language models: challenges for Critical Discourse Studies, Critical Discourse Studies, DOI: 10.1080/17405904.2024.2373733

To link to this article: https://doi.org/10.1080/17405904.2024.2373733

Published online: 11 Jul 2024.

CRITICAL DISCOURSE STUDIES
https://doi.org/10.1080/17405904.2024.2373733

The rise of large language models: challenges for Critical Discourse Studies

Mathew Gillings [a], Tobias Kohn [b] and Gerlinde Mautner [a]

[a]Institute for English Business Communication, Vienna University of Economics and Business, Vienna, Austria; [b]Department of Informatics, Karlsruhe Institute of Technology, Karlsruhe, Germany

ABSTRACT

Large language models (LLMs) such as ChatGPT are opening up new areas of research and teaching potential across a variety of domains. The purpose of the present conceptual paper is to map this new terrain from the point of view of Critical Discourse Studies (CDS). We demonstrate that the usage of LLMs raises concerns that definitely fall within the remit of CDS; among them, power and inequality. After an initial explanation of LLMs, we focus on three key areas of reflection. The first is a general stock-taking, where we look at CDS' theoretical underpinnings and what they imply for working with AI-generated language data. The second issue is authorship, where we assess the traceability of linguistic metadata and the ethically sensitive situation with regard to ownership of texts. The third area is linguistic homogenisation, where we examine how LLM usage privileges the mainstream. Afterwards, we explore ways in which LLMs could be used in research, and we discuss the implications of exploring their use in the classroom through a CDS lens. We close the paper with some observations on likely future developments in AI and how CDS can contribute with its distinctive theoretical, methodological and critical apparatus.

ARTICLE HISTORY
Received 30 September 2023
Accepted 25 June 2024

KEYWORDS
Artificial intelligence (AI); large language models (LLMs); authorship; ethics; power

1. Introduction

Since late 2022, large language models (LLMs) such as ChatGPT have been the talk of the town. Developed by the American software company OpenAI,[1] the *Chat Generative Pre-trained Transformer*, as its full name reads, is 'an artificial intelligence (AI) chatbot that processes and generates natural language text, offering human-like responses to a wide range of questions and prompts' (Doshi et al., 2023, p. 6). In the media, ChatGPT and similar software are being met with both alarm and grudging admiration. They are hailed for their potential to solve old problems – ranging from climate change to hitherto incurable diseases, and even saving a dog's life (e.g. The Economic Times, 2023). And they are condemned for creating new ones, for example by making certain jobs redundant (e.g. Mitchell, 2023), or being a potential threat to human rights (Rodrigues, 2020).

CONTACT Mathew Gillings mathew.gillings@wu.ac.at Institute for English Business Communication, Building D2, Vienna University of Economics and Business, Welthandelsplatz 1, 1020, Vienna, Austria @mathewgillings

In everyday private interaction too, ChatGPT and programs like it have become a constant presence. That ubiquity alone, and the socio-political relevance it entails, flags up AI as a prime candidate for Critical Discourse Studies (CDS) inquiry.

At the time of writing, in early 2024, ChatGPT is the best-known example of an LLM, but news of competing products are also making headlines. It is impossible to say at this stage whether, like *hoover* and *google* before it, *ChatGPT* will morph into a generic noun (or even verb). This is why we have adopted the safer terminological policy of using the non-branded term *large language model*, to refer to ChatGPT and similar systems applying AI.

Arguably, it is rare for new software applications to make such waves so quickly outside the computer-science community. The stakeholders in AI-related matters are many and varied, as are their concerns and agendas. To date, one of the most vocal groups seems to be educators worried about students using AI to cheat, but it is becoming clear that there is hardly a social domain that is not affected. Across academia, LLMs have piqued researchers' interest, and linguists are no exception (e.g. Crosthwaite & Baisa, 2023; Curry et al., 2024; Kohnke et al., 2023; Lin, 2023; Yu et al., 2024). After all, the use of AI has a wide range of linguistic applications and implications that potentially open up new research avenues.

Although this is quite a new research arena, critical discourse analysts are on their home turf. Many of the issues that we explore in this paper have also been discussed within the context of the rise of algorithms and the curation of other forms of representational content. The study of social media, for example, has been ripe for CDS input (see, for example, the range of papers published in Eposito & KhosraviNik, 2024). In future, we might consider LLMs to be the next step in an upward trajectory of more and more algorithmic interference in our social practice. Think about the leap from simple online shopping to recommender systems predicting your next purchase; then another leap to systems such as TikTok whose content is fully tailored to the user, based on choices they made in the past (Jungherr & Schroeder, 2023).

The main difference, between then and now, is that now we have a new type of language data, produced by machines rather than humans. This needs to be assessed critically, both in terms of its production, reception and social impact; sources need to be questioned; taken-for-granted assumptions challenged; ethical implications assessed; dominant discourses critiqued; and relations between language and society reappraised. CDS still has its primary aim 'to advance our understanding of how discourse figures in social processes, social structures and social change' (Flowerdew & Richardson, 2018, p. 1). Discourse is regarded as manifesting social reality, and social reality as shaping discourse (Fairclough, 1992a, p. 64). Those fundamentals have remained the same. However, with LLMs now on the scene, what has changed is the nature and status of the textual evidence before us. LLMs aren't simply another source of data. Thus, while the research terrain may look familiar, it is actually unchartered and potentially treacherous. This is why CDS cannot expect to continue doing 'business as usual'.

Our general aim in this paper is to reflect critically on what the arrival of LLMs could mean for CDS. Their impact is likely to be felt in three connected areas: (1) LLM output as a form of linguistic data; (2) LLMs as tools to help in the (linguistic) research process; and (3) AI in wider society, including the discursive 'fallout' of LLM applications in a range of social domains. Our paper is theoretical in nature and thus complementary to empirical papers that evaluate AI's potential to aid in CDS research (e.g. Curry et al., 2024).

Separately, but even more so in tandem, both empirical studies and conceptual pieces like ours should provide guidance to readers whose CDS projects touch on LLMs in some way or other. In addition to being reflective, our paper also has a normative and activist streak (again very much in the CDA/CDS tradition, as laid out, for example, by van Dijk, 1993). That is, we not only critique LLMs but also warn against naïvely embracing the promises they seem to hold.

The paper is organised as follows. We begin in Section 2 by providing some background, describing what LLMs are, how they work, and how it comes to be that its output ends up sounding so 'natural'. In Section 3, we move on to discuss LLMs within the context of traditional principles of CDS, focusing on issues such as authorship, power, and linguistic homogenisation. In Section 4, we explore areas where LLMs and academic research may intersect. After all, these language models undoubtedly have the potential to be useful in the research process, yet they must not be used uncritically simply because they are there. And in Section 5, we discuss the implications that LLMs are having on education. We conclude by summarising our observations and reflecting on what they mean for future research agendas and designs.

Here, early on, we acknowledge that in an area as vast and complex as AI, a paper-length treatise is bound to have gaps by both accident and design. Four of the latter we would like to flag up straight away. First, for reasons of focus as much as space, we will be concentrating on AI systems that process *text* and leaving to one side those that deal with images. Second, our remit in this paper does not include dealing with any of the technical issues that may arise when LLM-related language is prepared for linguistic analysis. Third, we deliberately steer clear of advocating concrete research designs and research questions. Our aim is to map the territory, not recommend specific routes through it. And finally, the field is moving at such a frenetic pace that we have also been shooting at a moving target in terms of the literature we could reasonably hope to catch up with and accommodate in our own piece. We have done our best to include papers that appeared between our first and revised submissions, but are very much aware that we will have missed many others.

The questions discussed in this paper are located at the interface of linguistics and computer science, and to reflect this, our author team consists of a computer scientist, who specialises in Computer Science Education (see Kohn, 2019; Phung et al., 2023) as well as two linguists, with specific experience in (corpus-assisted) discourse studies (see Hardt-Mautner, 1995; Gillings et al., 2023; Gillings & Mautner, 2024; Mautner, 2007). Our own backgrounds inevitably affect how we approach the topic and frame our argument, and as such our primary audience are researchers interested in working with large textual datasets.

Yet the composition of this author team and its preferred perspective should not be taken to imply that other combinations would not be fruitful or interesting. After all, it is unlikely there will remain a discipline which does not interact with the topic in some way. Work already published includes the areas of business, management and organisational studies (e.g. Paul et al., 2023), sociology (e.g. Balmer, 2023), marketing (e.g. Huang & Rust, 2021), and education (e.g. Tlili et al., 2023). Although we are unable in this paper to engage with the entire gamut of cross-disciplinary interactions, we hope that our reflections will prove relevant beyond the connection between computer science and linguistics.

Finally, another caveat. Given the novelty of the tool, and the speed at which IT is progressing, our comments are bound to be somewhat speculative, and history will probably prove us wrong on more than one count. Yet in writing up our thoughts, we made a conscious decision to allow ourselves rather more crystal-ball gazing than would normally be the case in an academic journal. We felt that under the circumstances a bolder approach was necessary to stay ahead of the game. For if everyone reserved judgement on LLMs until detailed and substantial empirical evidence was available, the research community would find themselves reduced to mere observers of, and commentators on, linguistic and social developments already underway, rather than taking an active part in shaping them. The point of no return is not a good place to be for researchers committed to having an impact on social reality. In keeping with CDS' activist credentials, most scholars in the field would rather not forego opportunities for mitigating or indeed preventing LLMs' more egregious repercussions. The aforementioned activist commitment is not well served if we are only ever wise *after* the event.

2. Background: what is a large language model (LLM)?

In order to appreciate the opportunities and constraints of LLMs, we first need to understand how they work. Essentially, what an LLM does is to take the beginning of a text, and then predict what the next word of that text might be, eventually building up phrases, sentences, and even complete paragraphs. This process can be used to generate text; in a chatbot, for example, a cycle of such predictions is triggered by a user-provided prompt (i.e. the question that the user 'asks' ChatGPT).

In order for such a language model to be successful, it must overcome two obstacles. First, it needs to extract those parts of the existing text that are salient in the sense that they are actually relevant for the prediction of the next word. Second, the model must account for the fact that the amount of text that is available for training the model is invariably limited, especially in comparison to the vast number of potential ways that words can be combined creatively; the model generalises from that. Both these challenges are naturally connected in that identifying relevant parts of text helps with generalisation by ignoring minor variations and thus highlighting more general patterns.

A key question in this context is how LLMs deal with 'meaning'. The simple answer is that they do not: computers generally work only on a syntactic level with no understanding of semantics whatsoever (cf. Bender & Koller, 2020). What a sophisticated LLM does is to produce meaningful text on the basis of syntax and probability. It learns from the vast number of examples in its training data how a particular word is slotted into a sentence, and what other words are likely to follow. Technically speaking, it achieves this by replacing words with numeric vectors that encode the usage patterns of those words. This so-called 'word embedding' by itself allows the computer to discover certain relationships: for example, the vectors for 'dog' and 'puppy' differ almost exactly the same way as those for 'cat' and 'kitten'.

Laws such as this form the basis of how machine learning works. An assumption is made that it is impossible (or impractical) to fully model the true underlying distribution of words; hence the attempt to approximate that distribution as well as possible via other routes. A good comparison to make is perhaps with Zipf's law: It gives us a good approximate distribution of the frequency of words, but it does not actually tell us the true and actual frequency. Thus, when interacting with ChatGPT and similar systems, we merely access an approximate statistic that has been derived from that initial dataset.

As we move from words to phrases, recognising patterns computationally becomes much more difficult. The success of modern LLMs stems from breakthroughs in exactly this area, provided by so-called 'transformer' architecture. A 'transformer' is a neural net that recognises word patterns in phrases, through the use of a mechanism called 'attention' (Vaswani et al., 2017). For any given text, the attention mechanism will tell you which words are most likely to influence the choice of the next word and which words are virtually irrelevant. ChatGPT's 'attention span' includes up to about two thousand words[2] from which it can pick the words most likely to predict the next (Brown et al., 2020). Anything outside its 'attention span', however, will be ignored entirely, which accounts for its tendency to 'forget' things that appeared earlier in a 'conversation'. Nevertheless, what in the end allows the system to generalise from samples to more abstract patterns is the combination of word embeddings and the attention mechanism.

The algorithmic operations involved are based on statistical distribution, and that inevitably peaks around the most obvious choices; hence, there is usually only a very small number of highly likely next words for any given phrase. A key question, then, is how you can prevent the generated texts from sounding too mechanical because they are based entirely on the most predictable choices. In order to create more variety, we may 'melt' some of the peak off and distribute the probabilities a bit more equally. This is controlled through a parameter known as 'temperature'. Increasing the temperature is akin to placing less emphasis on the actual frequency of words in the corpus and making more random choices.

At the risk of labouring the obvious, it is worth reminding ourselves that LLMs do not encode knowledge as such, but only linguistic data. An LLM is capable of completing the sentence 'The weight of an elephant is … ', only if this phrase has occurred in the corpus often enough that the correct answer points reliably to the 'most likely word'. Otherwise, the model will merely infer from the structure a number of expected answers and then pick one of those at random. It follows that texts may be linguistically sound (i.e. conforming to the expected patterns, as encountered in the training data), but still inaccurate in factual terms. When this happens, we speak of 'hallucination' – a phenomenon widely commented on, not least in the context of manufactured source references in academic writing (Walters & Wilder, 2023). Although this effect can be mitigated (for instance, by combining an LLM with an internet search engine), the primacy of linguistic fluency over factual accuracy is at the very core of LLM design. Here we have a potential tension between language and 'reality' – and unpacking that is something else that will have obvious appeal for critical discourse analysts.

Despite its success in overcoming the main obstacles of word prediction, the attention mechanism is still computationally expensive and requires a large number of text samples to find useful patterns. On the one hand, this means that only 'Big Tech' companies readily have the means to build the necessary computing infrastructure (Whittaker, 2021). On the other hand, there is a strong desire to reuse the attention mechanism once it has been trained; this is what the name 'generative pre-trained transformer' (GPT) refers to. The idea is that once trained, a transformer system can subsequently be adapted to new applications with a relatively moderate amount of effort and little additional training data.

Think of it this way: once you understand the general patterns of the English language, it is much easier to concentrate on the specific patterns used by a particular novelist.

Another key issue relates to the fine-tuning of the model. The full language model is trained in two stages: First, as many textual samples as possible are used to train its general capacity for pattern recognition. (The vast majority of these samples are in English, an issue that we will return to later.) Second, the model is fine-tuned by (statistically) strengthening desirable patterns and weakening undesirable ones. Unfortunately, fine-tuning may also introduce unwanted side effects, as evidenced by ChatGPT's worsening performance over time (Chen et al., 2023).

Fine-tuning is labour-intensive. It has to be carried out by humans, or at least with the aid of datasets that have been annotated by humans. Yet such fine-tuning is necessary, in large part, because of the highly problematic patterns that can be found in the underlying datasets – be they racist, sexist or violent in other ways. As investigative journalism revealed (Perrigo, 2023), in OpenAI's case, workers from Kenya were tasked with annotating problematic and often gruesome text samples to enhance the 'safety' of ChatGPT. These workers were employed by a San Francisco-based company, and severely underpaid by US standards. What is often hailed as a technological marvel for the benefit of the Global North, thus builds in large part on human labour from the Global South.

3. LLM output, natural language, and key principles of CDS

Artificial intelligence and computer-generated texts have been around for quite some time, yet the advent of modern LLMs in many ways now feels like a gamechanger. LLM output *looks* deceptively like language produced by humans, but on closer inspection, and based on what we saw in Section 2, it is not. So what is the nature of the beast that we are trying to capture analytically? This is a question of ontology. What *is* language generated by a machine? What kind of discourse do LLMs produce? And on a more mundane level, why should this be of concern to researchers generally, and CDS scholars specifically?

Potentially, LLMs have immediate attraction for all linguists who are short of time and resources, yet hungry for data. Increasingly, these can also be found among discourse analysts, as corpus-assisted discourse studies, or *CADS* for short, has been gaining momentum (Gillings et al., 2023), allowing qualitative researchers to work with larger, more representative datasets. Intrigued by LLMs, some may harbour hopes that time-consuming forms of data collection are now a thing of the past, and that instead, LLMs could be asked to produce huge datasets at lightning speed.

However, at least for the time being, that hope looks set to be dashed, and for reasons running deeper than the current state of technological development. Much more fundamental hurdles emerge if we examine the core principles of CDS (as first laid out in van Dijk, 1993 and summarised more recently in Mautner, 2010 for example). While many of these principles can be applied regardless of the data under analysis, two in particular seem to be at odds with how LLMs work and what texts they produce. First, CDS typically studies discourses through *naturally* occurring text and talk. Language produced by a machine does not quite seem to fit the bill. Second, CDS interprets text and talk by drawing on contextual knowledge. LLM-generated texts are not tethered to a context, at least not in a traditional sense. Let us examine these two principles in more detail.

The first begs the question of whether we should perhaps reappraise the dichotomy between natural and artificial data. Traditionally, most linguists would consider data 'natural' if it had not been elicited specifically for the purpose of analysis. Yet the advent of LLMs has complicated matters considerably. LLMs are trained on naturally occurring human text, but their output in response to queries is essentially a statistical, and thus 'artificial' product. Does this make LLM-generated text *fully* artificial? Or does it retain some natural quality because the 'raw material' it draws on was, after all, produced by humans without elicitation? The jury is still out on which category we should place LLM-produced text within, or indeed whether retaining the natural/artificial dichotomy makes sense at all. For the time being, it seems that LLM-generated language is perhaps best considered to be in a league of its own, not only sharing qualities from each, but representing a type of discourse *sui generis*.

The second point, about the degree to which LLM-produced text can be tied to context, also requires an open mind. Again, from a traditional perspective, we would argue that when texts are produced by an LLM, we can no longer clearly identify the ties between the AI-generated texts and their original extralinguistic environment (i.e. the context they were embedded in in the training data). LLM texts have no natural hinterland that can be explored and drawn upon as a sense-making device. However, if we interpret context more widely, to refer to any factor that is technically outside language but has a bearing on it, then LLM-generated output, too, relies on 'context': the underlying algorithms, the prompts (both individually and in terms of their sequence and history) as well as the LLM's fine-tuning, the training data, and the authors represented within it. Taken together, all these contextual factors have to be factored in when LLM texts are interpreted.

Among these, the question of authorship stands out as particularly important. Traditionally, authorship is associated with one person or a group producing texts, and negotiating meaning in and through interaction. LLMs wreak havoc with these notions. Of course, authors are still involved at some point; after all, they have produced the training data that LLMs feed on. Yet the software has created an amalgam from which individual source texts can no longer be extracted, and their authors remain in the shadows. Metaphorically speaking, a traditional dataset used in CDS is a collection of (ideally, clearly labelled) items in a storage cupboard; the output produced by an LLM, on the other hand, comes from a food blender.

To highlight the differences between data produced with and without the aid of AI, let us look at a hypothetical classic study of, say, representations of poverty in newspaper discourse (similar to the work of Paterson & Gregory, 2019, for example). In compiling and marking up our data, we are likely to want to consider newspapers of different political persuasions. We are also likely to identify different journalistic genres – such as leader articles and reports, and decide which ones to include or exclude. Some genres will have named authors, and those that do not can also be assumed to be aligned with the voice of the newspaper concerned. And as a final layer of accountability and influence, there are the paper's proprietors, who may or may not be known to shape editorial policy. All this information about authorship can be drawn on when it comes to interpreting the (corpus-based) findings, linking the emerging patterns to who uses them. Provided the dataset has been properly annotated, relevant metadata about sources, writers and speakers will be available, ready to be accessed as and when necessary. By contrast, when the textual output has been generated by an LLM, no information is available about authors, nor their social and discursive power. In fact, the very concept of authorship no longer makes sense. This does not necessarily render such outputs useless, but it does mean that any claims made about the findings must be qualified to reflect the opaque and elusive origins of the material.

There are several other effects of LLMs that are related to power. One is that their output will become more homogenous over time, with implications both on a technical and a social level. On the technical level, the key issue is this: as LLMs evolve from one version to another, the training data themselves will be 'deeply blended' (Guo et al., 2023) – that is, they will contain an increasing proportion of text that has itself been produced by an LLM or combined with an LLM output. This is problematic from a computer science point of view, as there are ripple effects from reusing the same data. Because those outputs are based on the most frequent patterns (with the appropriate 'temperature' control discussed in Section 2), outputs are relatively similar (though not identical) each time they are prompted. Yet examples of stylistic creativity – an unusual metaphor or colourful turn of phrase – are typically not found in its output, thus leading to a fairly banal standard (Padmakumar & He, 2023). All the less common lexical phenomena, which make a text stand out as distinctive, simply do not make it over the frequency threshold. Reusing more standard output, again and again, essentially leads to 'model collapse'; that is, the LLM trains on its own language and degenerates as a result (Shumailov et al., 2023). Effectively, the linguistic 'gene pool' becomes progressively more limited, and at present there appear to be no checks and balances to prevent this.

On a social level, we find that the more homogenised language becomes, the less weight is given to underrepresented groups. Social élites and mainstream views simply leave a significantly larger discursive footprint, and this advantage is perpetuated by AI. Given the nature of the algorithms involved, systems being trained predominantly on texts written by privileged author and speaker groups will inevitably produce more and more texts with the same characteristics. By the same token, dissenting voices risk being marginalised further, 'with non-conforming content pushed out of sight' (Jungherr & Schroeder, 2023, p. 5). Thus, existing biases and discrimination may be amplified, and hegemonic views reinforced (Jungherr, 2023, p. 7; Bender et al., 2021, p. 613). For example, when Atari et al. (2023) asked ChatGPT to complete the World Values Survey, and compared the results to those from across the world, they found that they were most similar to those responses from the USA. Arnold et al. (2018, p. 33) argue that there is a 'chain of bias', where 'biases in training data cause biases in system behavior, which [when used to assist humans in writing text] in turn cause biased-human generated products'. There is thus a real danger that LLMs reinforce racism (Adib-Moghaddam, 2023) and in fact resemble 'the algorithms of oppression' identified in search engines (Noble, 2018). In an interview quoted in Weil (2023), Bender elaborates on this in starker terms that will surely send alarm bells ringing for discourse analysts:

> The training data for ChatGPT is believed to include most or all of Wikipedia, pages linked from Reddit, a billion words grabbed off the internet. […] The humans who wrote all those words online overrepresent white people. They overrepresent men. They overrepresent wealth. What's more, we all know what's out there on the internet: vast swamps of racism, sexism, homophobia, Islamophobia, neo-Nazism. (Weil, 2023)

Another side-effect of how the training data is collected concerns the dominance of English. At risk of stating the obvious, given that English is by far the most common language found on the web, and given that LLMs pull texts from the web, it follows that English is the main language to be found within LLMs' training data. In the case of ChatGPT, the initial training set comprised 500 billion tokens, with 81% of them from internet-based datasets, 16% from books, and 3% from Wikipedia. In terms of word count, 93% of the data is in English (Brown et al., 2020). And so, because of this built-in bias, the cycle of English dominance continues.

On a related, albeit slightly different note, there is also evidence of human language becoming more and more homogenised to suit the LLM that we are now interacting with. This is known as *prompt engineering*, where we phrase our input prompts in such a way as to get a better output result from the model. Think about how you word Google searches; it is unlikely that you type out full questions, but instead tweak your linguistic style to match what experience tells you will produce the best results. Perhaps more worryingly, there is now evidence that text suggestions from an LLM have the power to influence the opinion of the user, even beyond the writing of the current text (Jakesch et al., 2023). If that is the case, then users end up not only sounding like the chatbot, but believing what it says too.

4. LLMs as a research tool

Over the past year, we have seen various ideas about how LLMs can be employed to aid in the research process: both as an additional source of information, and as an additional analytical tool. Regarding the former, LLMs could be used for advice on research tasks, explaining jargon-heavy statistical techniques in simplified terms, rewording complex passages of text, helping with professional writing (e.g. Cardon et al., 2023), and so on. As for the latter, LLMs could be used as a tool to conduct analyses; essentially, deploying them to interpret linguistic data. Curry et al. (2024) explored the utility of just that; they used ChatGPT to replicate three previous corpus-assisted discourse analyses. They found that whilst ChatGPT was reasonably effective at semantically categorising keywords (that is, categorising groups of decontextualised words), it was poor at concordance analysis (which naturally relies on additional context to interpret), and also poor at function-to-form analysis. At times, they found that the system made false inferences and incorrectly quoted data, which was made worse by the system's analyses not being replicable (that is, the same command led to different outputs). Yu et al. (2024) are slightly more upbeat about its potential, finding that in annotating apology components (a form of pragma-discursive analysis), GPT-4 (the LLM behind Bing's chatbot) did so with an accuracy approaching that of a human coder.

Clearly, LLMs perform analytical jobs with varying degrees of success, depending on the specific research task and the commands used to instruct them. The key consideration, then, is that the tasks we give to the LLM are those that they have been trained to do well. Asking an LLM to perform a task more suitable for humans sets us up for disappointment; in the same way that asking a human to perform a task more suitable for computers would too. Humans and computers are designed to do different things, and we risk evaluating the 'intelligence' of one with tests originally designed to test the other.

Thus, with that proviso in mind, recent research shows that LLMs hold considerable promise for CDS. If anything, their potential is likely to increase further as the software becomes more sophisticated, and researchers more adept at using it. Even so, the critically reflexive mindset which – ideally – guides CDS work must remain on the alert throughout the research process, from the early stages when research questions are first formulated to the final stages when data are interpreted and wider conclusions drawn about how discourse and society impact each other in a specific setting. As with other methods, we need to ask, for a start, whether we chose a particular research project because it was socially relevant and then looked for a suitable method, or whether we chose that project because it fitted our preferred method.

Several decades of pre-LLM corpus-assisted discourse analysis suggest rather strongly that automation is fascinating per se and comes with a risk of crowding out other approaches and obscuring alternative perspectives on the data. It is tempting to merely pay lip-service to the combination of quantitative and qualitative angles while being so engrossed in the former that little time, energy and creativity – not to mention space in publications – are left for the latter. Arguably, the more powerful the software tool, the more compelling the results and the less incentive for the researcher to take a step back and critically assess what impact the tool they have used has had on the process and product of their research.

Yet that assessment is crucial, at least for two related reasons. First, because the methodological road not taken could have led to equally, or perhaps even more interesting results; or it could have helped to put the LLM-assisted results into perspective. Second, because it is easy to forget that LLMs, as we outlined above, are shaped by the training data and the value judgements that feed into the training process. To trust an LLM blindly effectively means delegating the underlying ethics of the analytical venture to an algorithm.

5. Implications for teaching and learning

In an era increasingly influenced by LLMs, what can CDS contribute to teaching and learning, both in the language classroom and beyond? In terms of authorship, a central – and transferable – learning outcome is critical language awareness. This has always been an important skill (Fairclough, 1992b), but its significance has no doubt been increased by the existence of artificially-created texts. Students must learn to appreciate the statistical origins of LLM output as well as the key differences between it and naturally-occurring language. As teachers, we should nurture in our students not only a healthy distrust of *what* chatbots write but also *how* they write it; in short, students need to develop 'AI literacy' (Cardon et al., 2023). Equipped with a solid understanding of how language works in an AI context, students will also be in a better position to sniff out falsehoods even when they are hidden between truths.

The expected homogenisation of language also has a wide range of pedagogical implications which at this stage are difficult to evaluate. The more LLM-generated language is out there, the more frequently L1 and L2 learners will encounter it, and the greater the chance that they will model their own language on it. At this stage, LLM training data is undocumented (other than what OpenAI makes available on its website, in the case of ChatGPT), and there are no large-scale studies yet of the linguistic qualities of its output. Anecdotally, it seems that chatbots produce predominantly Standard English. Whether increased exposure to, and the resulting imitation of, this variety is good or bad is in the eye of the beholder. If you believe that users of ostensibly less prestigious varieties will benefit socially from increased access to this language, you will consider LLMs to have empowering potential (increasing speakers' chances on the job market, for example). If on the other hand, you believe that standard varieties inevitably crush the vernacular, then LLMs achieve the very opposite, *dis*empowering language users, robbing them of a key means of expressing their social identity, and alienating them from their communities. In practice, these two beliefs are not as incommensurable as they sound. After all, the social impact of acquiring a standard variety was a mixed blessing long before AI arrived on the scene. These social issues apart, LLM-induced homogenisation could be a valuable talking point in the classroom. It can quickly lead to discussions about the fundamental nature of language as a system in which a careful balance is constantly struck between structural constraints and creative freedom.

What also lends itself to being discussed in the classroom are our reflections on access to knowledge. Again, there are positive and negative sides to the problem. On the one hand, much of the world's knowledge is currently difficult for marginalised groups to access, and this issue is likely to be exacerbated further if chatbots such as ChatGPT cease to be freely available. For the time being at least, LLMs do have the potential to open up learning opportunities to groups who would otherwise be barred from them, and whose opportunities are curtailed as a result. On the other hand, students should be made aware that the mere process of generating LLM output invariably privileges those social actors that are overrepresented in the public domain anyway. Voices that are already quite loud in the training data become even louder still (Jungherr & Schroeder, 2023, p. 5), and it takes a critical lens to see this.

So, *quo vadis* education? There is no shortage of reports about teachers being creative in their approach to integrating AI in the classroom, for example by treating the chatbot as a seminar participant, encouraging students to critically interact with it, finding flaws in its argument, and engaging in debate (Darics & van Poppel, 2023). Others have asked students to compare their academic writing with that produced by a chatbot, thereby opening up discussions about good writing practices (Almirall, 2023). Exercises such as this foster creativity, critical reflexivity, and academic integrity. After all, when graduates leave university and enter the world of work where such algorithms are the norm, they must do so equipped with the necessary set of skills to appropriately assess LLM output.

Somewhat ironically, AI appears to be reviving classic principles of education that have stood the test of time, and that should be underpinning all university teaching and learning: critical reflection on the one hand, and personal dialogue on the other. It is true that these principles have come under increasing pressure through administrative constraints, mass enrolment and short-termist views of employability. However, their relevance is being thrown into sharp relief by the advent of LLMs, precisely because the latter are not very good at them. With these pressures in mind, it is unlikely that a romanticised vision of education as a constant stream of enlightening Socratic dialogue will come to pass. Yet the underlying spirit is definitely worth cultivating so that we can help students develop into discerning sceptics.

6. Conclusion

Large language models are clearly having a major impact on a variety of social domains and their associated academic fields. The impact on discourse and CDS was our main concern. From that perspective, new questions and challenges arise on a number of levels. The more ubiquitous and the more sophisticated LLMs become, the more essential it is to look critically at how they work, what kind of discourse they produce, and how their presence impacts different social groups.

A key observation in this context is that LLM-generated output is not *natural* in the way that discourse analysts would traditionally understand the term. It is a hybrid, based on the human-produced text that goes into the training data, but eventually created by statistical algorithms. The training data, in turn, have been shown to be heavily dominated by, and skewed towards, values and viewpoints originating in the Global North and exerting world-wide influence. What is more, LLM texts lack the rich contexts that truly natural texts and interactions are embedded in. Rather like robots, algorithms are not authors with biographies, interests, group affiliations and identities – in short, they lack the social characteristics typically factored in by 'thick' critical analyses. And because we cannot tell whose voice we hear, power structures and relationships between social actors also remain obscured.

Another angle that we briefly explored was whether LLMs could be used as a research tool, and at what stage in the research process this would be helpful and appropriate. What evidence there is suggests, not too surprisingly, that LLMs are on the whole more promising during phases that involve relatively mechanical steps, but generally disappoint when the analysis becomes more complex, fuzzier, and requires reflection and critical distance.

Looking ahead, with AI developing at breakneck speed, how should CDS react? What other challenges and opportunities for research are likely to arise? Three areas of interest spring to mind. First, we need to face difficult questions around copyright and ethics, which have in fact already led to high-profile court cases (see e.g. Creamer, 2023). Closer to home, there is already discussion within linguistics about these issues in corpus building (e.g. BAAL, 2021; Collins, 2019; Lutzky, 2021). Yet at least, with few exceptions, corpus-building projects within academia tend to be for the furthering of knowledge. When commercial interests are at stake, however, the implications are very different and potentially more serious. To take but one example, OpenAI, the creator of the ChatGPT software, transitioned from a non-profit to an at least partially for-profit organisation (referred to as 'capped profit'[3]). As language data is being hoovered up by for-profit entities, what does this mean for intellectual ownership? What does it mean for the public domain of the many, if its content is appropriated for the private gain of the few?

Second, we believe that exploring the discourse *about* LLMs is important. Researchers have already begun to look at speakers' attributing social agency to chatbots, and the implications this has on trust (Heaton et al., 2024). But there is more to it than that. The discourse about LLMs is very much part of the more general discourse about AI and robotics. Key discussions that have emerged are around the future of work (e.g. West, 2018) and AI's ethical repercussions (e.g. Stahl & Eke, 2024). Human-machine interaction, too, has moved to a new level, now involving AI applications, such as microchipping, that interact directly with the human body (Barnhizer & Barnhizer, 2019; Guzman et al., 2023). In the workplace, this technology enables new forms of control: opportunities for employers, threats to workers – and a call to action for the critical discourse analyst.

Third, CDS also has an important job to do in studying AI's wider discursive fallout in a variety of scenarios that may ultimately change the very fabric of society – whenever AI plays a part in highly sensitive areas such as governance, policing and judicial decision-making (Barnhizer & Barnhizer, 2019; Jungherr, 2023; Kasy, 2023). The data and algorithms employed 'are not neutral but instead have values embedded in their design, use, and output' (Joyce et al., 2021, p. 3). Crucially, these values are generally not accounted for, and their effect on the public arena 'can happen unobserved' (Jungherr & Schroeder, 2023, p. 6). CDS can contribute to public debate about automated decision-making by critically examining the role that discourse plays in it. Who makes the rules that govern the decisions, and how can those decisions be appealed against (Kasy, 2023)? Power plays a crucial role and is, as ever, inextricably intertwined with discourse. Furthermore, it is because AI systems are trained on past data that they are prone to perpetuating past patterns of inequality and discrimination rather than promoting social change (Jungherr, 2023, p. 3; Jungherr & Schroeder, 2023, p. 5). Here, too, critical discourse scholars have their work cut out for them.

Increasingly, any attempt to actually contain the spread of LLMs looks futile. It seems wiser therefore to meet the resulting challenges head-on; not in the spirit of a resigned 'if you can't beat them, join them', but with a defiant 'if you want to join them, beat them'. Beat them, that is, with our unique intellectual armoury – by all means in partnership with machines but as their master rather than their servant. Accordingly, those analytical skills will be at a premium that ChatGPT and its ilk are not very good at: meta-level reflection, critical thinking and the kind of creative thought that is generated through human-to-human interaction. LLMs might be getting better at these things too, it is true, but surely humans are still capable of being one step ahead. If we were to allow ourselves a final stab at crystal-ball gazing, we can be fairly confident that AI will not silence the creativity and ingenuity of researchers.

CDS will continue to play an important role, critiquing the social and discursive practices that are affected by LLMs, asking who benefits and who loses out, and whose values and perspectives get to dominate the underlying algorithms. CDS' conceptual and empirical toolbox is well stocked and up to the task. Nonetheless, grappling with a phenomenon as multi-layered as LLMs is a major intellectual challenge. It is also an ethical imperative. If CDS honours its activist commitment in AI now, we might for once succeed in closing the stable door *before* the horse has bolted.

Notes

1. https://openai.com/ (accessed 29 January 2024)
2. In actual fact, LLMs do not work with the concept of a 'word' but instead with 'tokens', where longer words comprise several tokens and which would also include punctuation marks, for instance. ChatGPT's attention span of 2048 tokens therefore effectively translates into fewer words.
3. https://openai.com/blog/openai-lp

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributors

Mathew Gillings is an Assistant Professor at the Vienna University of Economics and Business. Before moving to Vienna, he completed his PhD in Linguistics at Lancaster University. His research interests include the studying the discourse of corporate wrongdoing, deception, and politeness in the workplace. He is a co-author of *Corpus-Assisted Discourse Studies* (CUP, 2023) and the author of *Corpus Linguistic Approaches to Deception Detection* (Routledge, 2024).

Tobias Kohn is Professor of Computer Science Education at Karlsruhe Institute of Technology, and a Quondam Fellow at Hughes Hall College, University of Cambridge. His primary research interest is in compilers, interpreters, programming languages, and education, with a particular focus on error detection in program code.

Gerlinde Mautner is Professor of English Business Communication at the Vienna University of Economics and Business. She pursues research interests located at the interface of language, society, business and the law. Her work also has a strong methodological focus, concerned in particular with the relationship between corpus linguistics and discourse studies.

ORCID

Mathew Gillings http://orcid.org/0000-0002-6114-0117
Tobias Kohn http://orcid.org/0000-0002-9251-8944
Gerlinde Mautner http://orcid.org/0000-0002-5529-3184

References

Adib-Moghaddam, A. (2023). *Is artificial intelligence racist? The ethics of AI and the future of humanity*. Bloomsbury.

Almirall, E. (2023, February 24). Three creative ways to use ChatGPT in class. *Times Higher Education*. https://www.timeshighereducation.com/campus/three-creative-ways-use-chatgpt-class

Arnold, K. C., Chauncey, K., & Gajos, K. Z. (2018). Sentiment bias in predictive text recommendations results in biased writing. *Proceedings of the 44th Graphics Interface Conference* (pp. 33–40).

Atari, M., Xue, M. J., Park, P. S., Blasi, D. E., & Henrich, J. (2023). Which humans? PsyArXiv preprint. https://doi.org/10.31234/osf.io/5b26t

BAAL (The British Association for Applied Linguistics). (2021). Recommendations on good practice in applied linguistics. https://www.baal.org.uk/wp-content/uploads/2021/03/BAAL-Good-Practice-Guidelines-2021.pdf

Balmer, A. (2023). A sociological conversation with ChatGPT about AI ethics, affect and reflexivity. *British Sociological Association.*, *57*(5), 1249–1258. https://doi.org/10.1177/00380385231169

Barnhizer, D., & Barnhizer, D. (2019). *The artificial intelligence contagion*. Clarity Press.

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 610–623).

Bender, E. M., & Koller, A. (2020). Climbing towards NLU: On meaning, form, and understanding in the age of data. In D. Jurafsky, J. Chai, N. Schluter, & J. Tetreault (Eds.), *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 5185–5198). Association for Computational Linguistics.

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., … Amodei, D. (2020). Language models are few-shot learners.

In H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, & H. Lin (Eds.), *Proceedings of the 34th International Conference on Neural Information Processing Systems* (pp. 1877–1901). Curran Associates Inc.

Cardon, P., Fleischmann, C., Aritz, J., Logemann, M., & Heidewald, J. (2023). The challenges and opportunities of AI-assisted writing: Developing AI literacy for the AI age. *Business and Professional Communication Quarterly*, *86*(3), 257–295. https://doi.org/10.1177/23294906231176517

Chen, L., Zaharia, M., & Zou, J. (2023). How is ChatGPT's behavior changing over time? arXiv preprint, arXiv:2307.09009. https://doi.org/10.48550/arXiv.2307.09009

Collins, L. (2019). *Corpus linguistics for online communication: A guide for research*. Routledge.

Creamer, E. (2023, July 05). Authors file a lawsuit against OpenAI for unlawfully 'ingesting' their books. *The Guardian*. https://www.theguardian.com/books/2023/jul/05/authors-file-a-lawsuit-against-openai-for-unlawfully-ingesting-their-books

Crosthwaite, P., & Baisa, V. (2023). Generative AI and the end of corpus-assisted data-driven learning? Not so fast!. *Applied Corpus Linguistics*. Advance online publication, *3*(3), https://doi.org/10.1016/j.acorp.2023.100066

Curry, N., Baker, P., & Brookes, G. (2024). Generative AI for corpus approaches to discourse studies: A critical evaluation of ChatGPT. *Applied Corpus Linguistics*, *4*(1), 100082. https://doi.org/10.1016/j.acorp.2023.100082

Darics, E., & van Poppel, L. (2023). Debate: ChatGPT offers unseen opportunities to sharpen students' critical skills. *The Conversation*. https://theconversation.com/debate-chatgpt-offers-unseen-opportunities-to-sharpen-students-critical-skills-199264

Doshi, R. H., Bajaj, S. S., & Krumholz, H. M. (2023). ChatGPT: Temptations of progress. *The American Journal of Bioethics*, *23*(4), 6–8. https://doi.org/10.1080/15265161.2023.2180110

The Economic Times. (2023, March, 28). ChatGPT saves dog's life, gives correct diagnosis that even vet could not identify. *The Economic Times*. https://economictimes.indiatimes.com/news/new-updates/chatgpt-saves-dogs-life-gives-correct-diagnosis-that-even-vet-could-not-identify/articleshow/99053583.cms

Eposito, E., & KhosraviNik, M. (2024). *Discourse in the digital age: Social media, power, and society*. Routledge.

Fairclough, N. (1992a). *Discourse and social change*. Polity Press.

Fairclough, N. (1992b). *Critical language awareness*. Routledge.

Flowerdew, J., & Richardson, J. E. (2018). Introduction. In J. Flowerdew & J. E. Richardson (Eds.), *The Routledge handbook of critical discourse studies* (pp. 1–10). Routledge.

Gillings, M., & Mautner, G. (2024). Concordancing for CADS: Practical challenges and theoretical implications. *International Journal of Corpus Linguistics*, *29*(1), 34–58. https://doi.org/10.1075/ijcl.21168.gil

Gillings, M., Mautner, G., & Baker, P. (2023). *Corpus-assisted discourse studies*. Cambridge University Press.

Guo, Y., Shang, G., Vazirgiannis, M., & Clavel, C. (2023). The curious decline of linguistic diversity: Training language models on synthetic text. arXiv preprint, arXiv:2311.09807. https://doi.org/10.48550/arXiv.2311.09807

Guzman, A. L., McEwan, R., & Jones, S. (Eds). (2023). *The SAGE handbook of human-machine communication*. SAGE.

Hardt-Mautner, G. (1995). 'Only connect': Critical discourse analysis and corpus linguistics. UCREL Technical Paper 6, Lancaster University. http://ucrel.lancs.ac.uk/papers/techpaper/vol6.pdf

Heaton, D., Nichele, E., Clos, J., & Fischer, J. E. (2024). "ChatGPT says no": Agency, trust, and blame in Twitter discourses after the launch of ChatGPT. *AI and Ethics*, https://doi.org/10.1007/s43681-023-00414-1

Huang, M.-H., & Rust, R. T. (2021). A strategic framework for artificial intelligence in marketing. *Journal of the Academy of Marketing Science*, *49*(1), 30–50. https://doi.org/10.1007/s11747-020-00749-9

Jakesch, M., Bhat, A., Buschek, D., Zalmanson, L., & Naaman, M. (2023). Co-writing with opinionated language models affects users' views. In A. Schmidt, K. Väänänen, T. Goyal, P. O. Kristensson, A.

Peters, S. Mueller, J. R. Williamson, & M. L. Wilson (Eds.), *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery.

Joyce, K., Smith-Doerr, L., Alegria, S., Bell, S., Cruz, T., Hoffman, S. G., Noble, S. U., & Shestakofsky, B. (2021). Toward a sociology of artificial intelligence: A call for research on inequalities and structural change. *Socius: Sociological Research for a Dynamic World*, *7*, 237802312199958–11. https://doi.org/10.1177/2378023121999581

Jungherr, A. (2023). Artificial intelligence and democracy: A conceptual framework. *Social Media + Society*, *9*(3), 1–14. https://doi.org/10.1177/20563051231186353

Jungherr, A., & Schroeder, R. (2023). Artificial intelligence and the public arena. *Communication Theory*, *33*(2-3), 164–173. https://doi.org/10.1093/ct/qtad006

Kasy, M. (2023). *The political economy of AI: Towards democratic control of the means of prediction*. INET Oxford Working Paper No. 2023-06. Institute for New Economic Thinking at the Oxford Martin School.

Kohn, T. (2019). The error behind the message: Finding the cause of error messages in python. In Proceedings of the 2019 ACM SIGCSE Technical Symposium on Computer Science Education (SIGCSE '19). (pp. 524-530). Association for Computing Machinery.

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal*, *54*(2), 537–550. https://doi.org/10.1177/00336882231162868

Lin, P. (2023). ChatGPT: Friend or foe (to corpus linguists)? *Applied Corpus Linguistics*. Advance online publication, *3*(3), https://doi.org/10.1016/j.acorp.2023.100065

Lutzky, U. (2021). *The discourse of customer service tweets*. Bloomsbury.

Mautner, G. (2007). Mining large corpora for social information: The case of elderly. *Language in Society*, *36*(1), 51–72. https://doi.org/10.1017/S0047404507070030

Mautner, G. (2010). *Language and the market society: Critical reflections on discourse and dominance*. Routledge.

Mitchell, A. (2023, January, 25). ChatGPT could make these jobs obsolete: 'The wolf is at the door'. *New York Post*. https://nypost.com/2023/01/25/chat-gpt-could-make-these-jobs-obsolete/

Noble, S. (2018). *Algorithms of oppression: How search engines reinforce racism*. NYU Press.

Padmakumar, V., & He, H. (2023). Does writing with language models reduce content diversity? arXiv preprint, arXiv:2309.05196. https://doi.org/10.48550/arXiv.2309.05196

Paterson, L. L., & Gregory, I. N. (2019). *Representations of poverty and place: Using geographical text analysis to understand discourse*. Palgrave.

Paul, J., Ueno, A., & Dennis, C. (2023). ChatGPT and consumers: Benefits, pitfalls and future research agenda. *International Journal of Consumer Studies*, *47*(4), 1213–1225. https://doi.org/10.1111/ijcs.12928

Perrigo, B. (2023, January, 18). Exclusive: OpenAI used Kenyan workers on less than $2 per hour to make ChatGPT less toxic. *TIME Magazine*. https://time.com/6247678/openai-chatgpt-kenya-workers/

Phung, T., Padurea, V.-T., Cambronero, J., Gulwani, S., Kohn, T., Majumdar, R., Singla, A., & Soares, G. (2023). Generative AI for programming education: Benchmarking ChatGPT, GPT-4, and human tutors. *ICER*, *2*, 41–44. https://doi.org/10.48550/arXiv.2306.17156

Rodrigues, R. (2020). Legal and human rights issues of AI: Gaps, challenges and vulnerabilities. *Journal of Responsible Technology*, *4*, 100005. https://doi.org/10.1016/j.jrt.2020.100005

Shumailov, I., Shumaylov, Z., Zhao, Y., Gal, Y., Papernot, N., & Anderson, R. (2023). The curse of recursion: Training on generated data makes models forget. arXiv preprint, arxiv:2305.17493. https://doi.org/10.48550/arXiv.2305.17493.

Stahl, B. C., & Eke, D. (2024). The ethics of ChatGPT – Exploring the ethical issues of an emerging technology. *International Journal of Information Management*, *74*, 102700. https://doi.org/10.1016/j.ijinfomgt.2023.102700

Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, *10*(1), 15. https://doi.org/10.1186/s40561-023-00237-x

van Dijk, T. A. (1993). Principles of critical discourse analysis. *Discourse & Society*, *4*(2), 249–283. https://doi.org/10.1177/0957926593004002006

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, & R. Garnett (Eds.), *Advances in neural information processing systems 30* (pp. 5998–6008). https://papers.nips.cc/paper_files/paper/2017

Walters, W. H., & Wilder, E. I. (2023). Fabrication and errors in the bibliographic citations generated by ChatGPT. *Scientific Reports*, *13*(1), 14045. https://doi.org/10.1038/s41598-023-41032-5

Weil, E. (2023, March, 01). You are not a parrot. *Intelligencer*. https://nymag.com/intelligencer/article/ai-artificial-intelligence-chatbots-emily-m-bender.html

West, D. M. (2018). *The future of work: Robots, AI, and automation*. Brookings Institution Press.

Whittaker, M. (2021). The steep cost of capture. *Interactions*, *28*(6), 50–55. https://doi.org/10.1145/3488666

Yu, D., Li, L., Su, H., & Fuoli, M. (2024). Assessing the potential of LLM-assisted annotation for corpus-based pragmatics and discourse analysis: The case of apologies. *International Journal of Corpus Linguistics*.