

Bettina Migge*, Britta Schneider and Iker Erdocia

Sociolinguists against the AI hype? AI tools in sociolinguistics scholarship

<https://doi.org/10.1515/lingvan-2025-0212>

Received October 17, 2025; accepted April 9, 2026; published online September 14, 2026

Abstract: This paper examines sociolinguists' views of and engagement with artificial intelligence technology in sociolinguistics scholarship, an area that has received little systematic attention to date. Drawing on a global survey of 274 respondents, the study reveals that while automatization and AI tools are mainly used for efficiency and experimentation purposes, their adoption for data collection and analysis remains limited. Moreover, sociolinguists express strong reservations about AI, including concerns over ethical, environmental, research integrity, and power-related issues. Our data show that AI technology has not disrupted the epistemological or methodological foundations of sociolinguistics research. We argue that under structural pressures to maximize productivity and efficiency, sociolinguistics positions itself as a cautious field, calling for critical reflection on the so-called AI hype. The study urges continued scrutiny of AI's implications for the production of sociolinguistic knowledge.

Keywords: sociolinguistics; research ethics; big tech; AI and linguistics

1 Sociolinguistics and the role of AI tools in research

Sociolinguistics examines the relationship between language and society, addressing questions of language, identity, and power. It has a strong focus on empirical research and a developing range of methodological approaches to study how language is influenced by and affects social factors, social structure, concepts of language, governance, and linguistic behaviour (Meyerhoff 2011). While investigation of digital language practices has become more popular recently (Androutsopoulos 2016), reflexive inspection of technology, including AI and its impact on language practices, started somewhat late.¹ In this article, we discuss the results of a survey aimed at understanding the sociolinguistic community's use of, and attitudes towards, AI language technologies in work and research contexts where the adoption of these technologies is relatively recent. Whereas there is an increasing interest in sociolinguistic perspectives in technological realms, our questionnaire is primarily interested in the "core" sociolinguistic community and thus in researchers whose prime aim is to understand the interaction between language and society, where technology has not been at the centre of attention (Lamoureux et al. 2026). Our study provides tentative indications about AI technologies' potential to reconfigure methods and theories in sociolinguistics.

Sociolinguistic research on digital language practices has identified various challenges. Maly (2022) and Kelly-Holmes (2022) demonstrate that online interactions, especially those moderated algorithmically, cannot be investigated in the same way as face-to-face interactions. Erdocia et al. (2026) show that private companies make decisions about language norms in digital space, and Kelly-Holmes (2024) discusses methodological struggles that emerge as we cannot be sure to what extent language data is co-produced by algorithmic technologies. Due to the proprietary nature of training data and algorithms, the reasons for specific AI outputs are difficult for researchers

¹ We use "AI tools" and "generative AI" interchangeably, defining them as computational systems that rely on automatic pattern recognition. The survey did not define them, as we were interested in participants' understandings.

***Corresponding author: Bettina Migge**, University College Dublin, Dublin, Ireland, E-mail: bettinamigge@ucd.ie. <https://orcid.org/0000-0002-3305-7113>

Britta Schneider, University of Vienna, Vienna, Austria. <https://orcid.org/0000-0003-0377-5808>

Iker Erdocia, Dublin City University, Dublin, Ireland. <https://orcid.org/0000-0003-2459-1346>

to assess, and it is not clear whether we can treat AI-generated letter and sound sequences as “language” (Erdocia et al. 2024). Moreover, linguistic biases in automatically generated language, be it chatbots, voice assistants, automatic transcription, or machine translation applications, have not been resolved (Koencke et al. 2020; Markl 2022; Savoldi et al. 2021). They cannot be easily ignored by sociolinguists as they hint at deeper social struggles and new power hierarchies that emerge in relation to the development and distribution of AI systems (Hao 2025) and may impact future language practices and language ideologies, leading to new forms of sociolinguistic hierarchies.

Some research navigates these challenges by using AI-driven technology in a supportive function only in conjunction with rigorous researcher-based assessment and for only some time-consuming tasks, such as transcription of spoken language data or to automatically extract, organize, or quantify data in the case of sociophonetic and forensic research (see other contributions in this special issue). Other research, also referred to as computational sociolinguistics (Perez-de-Arriluzea-Madariaga 2025), is interested in modelling language use and variation across social contexts and their relationship to social variables using large social media corpora and computational methods. Grieve et al. (2025) investigate the representation of language varieties using computational methods and urge greater attention to the sociolinguistic diversity and representativeness of input data to reduce social bias and enhance performance. Focusing on lexical phenomena, this research automatically extracts and correlates data with established social categories (e.g., age and class) with little researcher-based intervention, and results are not re-embedded in a social analysis (Nguyen et al. 2016). While modelling challenges are acknowledged and refinements are proposed (e.g., Grieve et al. 2025; Nguyen 2025; Nguyen et al. 2016), a strict correlational and “treasure-trove” (Kelly-Holmes 2022) analytical approach is pursued.² Due to an absence of open discussion, we know very little about how the wider (non-computational) sociolinguistics community reflects on AI tools, the challenges it has with them, and its ways of dealing with them or how it manages them.

For the purposes of this special issue, we examine this focusing on two broader areas:

1. To what extent are sociolinguists using AI tools in their academic work and, more particularly, when investigating language practices?
2. What are their perceptions of AI technologies, and how do these impact their AI usage and non-usage?

The investigation shows that while sociolinguists use AI technology in a supportive function for a broad range of editing processes, they, at least at the time when the survey was conducted (2025), make comparatively little and prudent use of it for core research activities.

2 Methodological approach – scoping review and survey design

As a first step, we initially followed Paulus et al.’s (2026) approach for assessing the use of AI in social science research. We manually inspected articles published in two major sociolinguistic journals (*Language in Society* and the *Journal of Sociolinguistics*) between 2024 and mid-2025 to see how researchers employ AI tools in their research.³ This did not prove very fruitful. Articles did not explicitly mention the use of AI. Of the 72 articles examined, four employed the statistical package R for quantitative analysis, one used a Python script for extracting data from a social media platform, another one used an algorithmic tool for modelling data, and two articles employed automatic tools for textual analysis. This shows that research automatization tools have a place in some sociolinguistic research but it was not possible to glean further details. This prompted us to conduct a questionnaire survey.

The survey, entitled “The role of AI in sociolinguistics research”, ran between June 2025 and September 2025, used Survey Monkey, and was distributed via major mailing lists and our professional networks using a snowball method. It aimed to inspect the prevalence of and reasons for AI use (including non-use), and the types of AI

² The “treasure trove” metaphor describes methods that analyse digital traces of human interactions with little reflection on how data come about and whose and which practices remain thereby invisible.

³ The *Journal of Sociolinguistics* published a short discussion piece and nine short responses on AI in 2024.

applications in our research community. It consisted of 26 open-ended and closed questions, devised on the basis of existing research on AI use among researchers in related disciplines (Paulus et al. 2026) and issues raised in sociolinguistic research (e.g., Grieve et al. 2025; Kelly-Holmes 2022, 2024; Maly 2024; Markl 2022), and in discussions with other researchers. The first three open-ended questions asked about the AI tools that participants currently use and plan to use to gauge their understanding of AI, and the languages in which they use them (Q2, Q3, Q4). The next set of questions probed participants' use of AI tools in their wider academic work (Q5, Q6), their use in their research practice (Q7, Q8, Q11, Q12), the type of research (Q9, Q10) and the reasons for using them (Q13, Q14). We also elicited people's views about AI (Q15, Q16) and how its use has affected conceptualizations of language (Q17, Q18), its relationship to society (Q19, Q20) and about ethical and other concerns (Q21, Q22, Q23). To avoid negative priming of respondents, we placed questions on concerns about AI at the end of the questionnaire. This was interpreted by some respondents as a sign of our positive attitudes towards AI. The final four questions elicited some basic demographic data such as the participant's employment status, the type of institutional context and broad region they work in, and their gender. With the exception of the first question, which asked participants to provide informed consent, participants could skip questions they did not want to answer.

In summary, the survey consists of a mix of closed and open questions that allowed participants to provide their own perspectives. Respondents made very active use of these open questions. Our analysis draws on the answers to both types of questions. Open questions are analysed using qualitative content analysis. Here, we focus on questions dealing with understandings of AI and how and why AI tools are used.⁴

3 How sociolinguists use (or do not use) AI

A total of 274 people completed the survey. Of these, about half were faculty (assistant to full professor), about a fourth were PhD students, and the remaining fourth were employed as postdocs or teaching fellows or in teaching and research-related posts. Nearly 99 % worked at an educational institution with only three people employed outside of this sector. Nearly two-thirds identified as women and just under a third as men, and the remaining participants either preferred not to disclose their gender or provided another response. More than two-thirds were based in Europe, followed by North America (10 %), Asia–Australia (9 %), and other parts of the world such as South and Central America (3 %), the Middle East (2 %), and Africa (1 %). While we strove to reach sociolinguists from across the globe, this geographical and linguistic imbalance needs to be kept in mind.

Instead of imposing our understanding of AI on survey participants, we first elicited respondents' understandings of AI by asking participants to name AI applications that they had used and were intending to use (Q2, Q3). Participants' views strongly converged in that 92 % of the 175 responses cited large language models (LLMs). OpenAI's ChatGPT was by far the most commonly mentioned LLM (90 %), while Gemini (20 %), Claude (8 %), Microsoft CoPilot (18 %), Perplexity (12 %), and Deepseek (3 %) were mentioned considerably less often. Some participants declared using several LLMs. Beyond LLMs, participants also cited using AI-powered translation software such as Google Translate and DeepL, writing aids such as Grammarly, and more specialized applications such as Elicit to scrape data from the internet, Whisper to transcribe texts, ResearchRabbit to help with literature reviews, Canva to design multimodal texts, and voice-to-text and image-to-text or text-to-image tools. Of those who did not list LLMs (8 %), 43 % reported not using any automatization tools, while 57 % listed machine translation software such as DeepL or Google Translate or transcription tools such as Whisper. This suggests that respondents have very similar understandings of AI as largely relating to LLMs and are aware that many research support tools are AI-driven. There were also questions about what counts as AI. Some respondents complained about our survey's lack of definition of AI, while others rightly noted that search engines are essentially "AI tools": "I guess we all use the AI that are used automatically on websites like Google etc" (Q2/#7).

Participants' strong awareness of AI and AI-driven tools did not lead to universal usage. In response to the question "How are you using AI technologies in your administrative and teaching activities?" (Q5), strikingly,

⁴ We refer to questions with its number in the survey (e.g., Q6); free answers are indicated by a hashtag and the comment number (e.g., Q6/#5).

more than a third of the participants stated not using AI tools for their administrative and teaching activities at all (Table 1). Those who use AI employ AI tools, most likely LLMs, in support functions to obtain ideas for activities, to prepare and proofread writing, and to draft standardized texts (e.g., emails, reference letters).

In the follow-up question, non-use was justified by strong ethical reservations, possibly prompted by our decision to place discussion of ethical reservations at the end of the questionnaire. One respondent argued that they “do not find any of the activities listed above [i.e., Table 1] taxing enough to justify the use of AI shortcuts” (Q6/#59). Those who said that they use or had used AI tools generally expressed reservations about their capacities. Several respondents explained in the follow-up question the need for human oversight: “There is always a considerable chunk of checking it manually” (Q6/#3); and one respondent even stated that AI technologies cannot (yet) be trusted “to write anything for me from scratch” (Q6/#65). Others, however, conceptualized them as a colleague or “a collaborator for these activities. I do not have AI just do them for me” (Q6/#12); or as a tool to help out: “It’s less useful for creating lesson plans; [...] But if I feel stuck, it sometimes gives me some ideas to get started again” (Q6/#4). Many respondents said that they use AI tools to check English grammar and spelling, for “use of foreign languages, including in emails, papers, and academic presentations” (Q6/#10), or to “shorten/reformulate/condense written texts, develop learning plans and timetables” (Q6/#36). Repetitive tasks such as “administrative work (class lists or groupings, calculate grades, etc.)” (Q6/#97) are also often delegated to AI but subsequently checked. Checking students’ use of them or to invite “students to engage critically with AI output” (Q6/#37) were also mentioned.

Even fewer respondents affirm using AI in their core research activities (Q7); see Table 2. Only 17 % of respondents indicate that they use AI tools for data analysis and 6 % for data collection. Its use is more common in a supportive role in the preparatory and final stages of research. About 18 % state using AI to generate ideas for research projects and their design, nearly 30 % to summarize writing, and 16 % to synthesize literature. As in the administrative and teaching realms, generative AI is widely employed for various editing tasks, ranging from enhancing grammar, structure, and stylistic fit of the researcher’s written outputs to developing illustrations.

In relation to *how* AI is used in research (Q8: “If you have used AI technology in your research, how have you used it?”), Table 3 strikingly shows that over 50 % do not use it, and almost one-third of participants have used it to critically investigate AI technologies’ capabilities. This suggests that many respondents take a prudent and tentative approach to AI technology, employing it comparatively little and mostly to evaluate its capabilities rather than systematically embedding it in the research process. While 20 % report using AI to analyse their own corpus or dataset that they had personally collected using surveys, interviews, or textual materials, others use AI to analyse data gathered from the web (10 %), to collect data (6 %), and to analyse corpora resulting from automated processes of extracting information (1 %). While these latter uses are common in computational sociolinguistics, they are little used among respondents, suggesting that most respondents probably come from

Table 1: Percentage of respondents who use AI tools in administration and teaching.

Use of AI in administration and teaching	%
Get ideas for lesson plans and learning activities for students	46 %
I do not use AI technologies in my administrative and teaching activities	39 %
Write emails to students and staff	30 %
Proof and enhance my PowerPoints and other learning materials	27 %
Expedite administrative work (class lists, groupings, calculate grades, etc.)	19 %
Summarize reading materials for meetings	19 %
Write reference letters for students	17 %
Create outreach materials	17 %
Write summative comments for student essays, projects, theses, etc.	12 %
Write peer reviews	9 %
Generate agendas for committee meetings	7 %
Check and grade my students’ (textual) assessment	5 %

Table 2: Percentage of respondents who use AI technology in research.

Core research activity where AI is used	%
Copy-edit academic texts (improve writing style, register, grammar, etc.)	51 %
Summarizing text (e.g., research articles, reports, research meetings)	29 %
Support writing academic texts, specifically to improve the structure	28 %
I do not use AI technologies for my research activities	27 %
Generate research ideas (e.g., research questions)	18 %
Generate ideas about how to design a research project	18 %
Analyse data	17 %
Prepare research projects/academic texts (e.g., lit reviews, synthesize literature)	16 %
Support writing of academic texts to improve the content	16 %
Create illustrations (e.g., figures) for my research	14 %
Create materials about research for non-academic audiences	14 %
Find and write funding applications	10 %
Collect data	6 %

non-computational traditions. It also suggests that question-order effects are unlikely, as respondents' negative attitudes towards AI do not appear to have been primed by the preceding questions.

In the following two open questions (Q9 and Q10, "Do you use AI technologies to enhance quantitative [Q9]/ qualitative [Q10] data analysis? ... If yes, for which steps in the analysis?"), participants could expand on their beliefs about and use of AI tools in core research activities, which respondents did extensively (102 and 94 responses, respectively). Most sociolinguists report not using AI (60 %) for quantitative analysis. Qualitative analysis of answers to corresponding open-ended questions shows that if used, these tools are employed for code generation and troubleshooting (e.g., writing, adapting, or debugging scripts in R or Python to streamline statistical work), data management and visualization (e.g., cleaning datasets and producing graphical outputs), and as a learning and support resource (e.g., clarifying analytical logic). An important theme that emerged is caution towards and mistrust of AI tools, with some participants explicitly avoiding AI due to concerns around its unreliability, its opacity, or ethics: "I have never used AI to analyze my data. I find that to be fraudulent at this point" (Q9/#83).

Responses to Q10 indicate an even more limited and cautious approach to AI, as 70 % engaging in qualitative research say that they are not using AI tools. Uses of AI include preparatory tasks, such as theme generation and coding suggestions, data organization, and management; and, to a lesser extent, tentative implementation of AI-assisted coding and annotation, with tools like ChatGPT, MAXQDA, or Atlas.ti. Once more, a recurring theme of ethical mistrust strongly emerges, with respondents emphasizing privacy, authenticity, and interpretive depth as reasons for avoiding AI. This theme is exemplified in the following three comments: "Hell no" (Q10/#59); "I think it would be an ethics violation to submit interview data to ChatGPT" (Q10/#5); "I don't believe in outsourcing any of my creative processes. I know I am old fashioned. I think there is still something important to be gained in terms of

Table 3: Percentage of respondents who use AI in particular ways in research.

Ways in which AI is used in research	%
None of the above	51 %
To critically investigate its capabilities	30 %
Analyse own corpus/dataset based on own collection (surveys, interviews, conversational language data collection, etc.)	20 %
Analyse own corpus based on data collected from the web and curated	10 %
Used AI to collect data for research	6 %
Analyse own corpus, which was created using web-crawling techniques	1 %

using my brain and expanding my own problem solving skills by doing everything myself. Even the tedious parts of research!” (Q10/#63).

Thus, while sociolinguists show an incipient pragmatism and some curiosity towards AI’s affordances, their stance reflects critical awareness of AI technologies’ limitations. Closer examination of the motivations for using AI technology (Q13) revealed that sociolinguists are mainly encouraged by the promise of efficiency and curiosity (Table 4). More than half of respondents (56 %) see it as speeding up time-consuming tasks, showing a pragmatic orientation towards optimizing research workflows. Again, exploratory uses of these technologies dominate, with 46 % using AI out of curiosity rather than as part of a clearly defined methodological or analytically driven purpose. This suggests that overall research approaches have not changed due to AI. A smaller but still relevant number of respondents mention other reasons, including 16 % acknowledging that they find AI methodologically or conceptually inspiring and 12 % being encouraged by colleagues or collaborators. Fewer respondents (7 %) report that AI has enabled them to identify new patterns undetectable through human analysis, and only a minority use it for large-scale data analysis (5 %) or due to institutional guidance (3 %). Interestingly, the representativeness of large datasets does not seem to be a factor for including AI tools in research.

Around one in four participants selected “None of the above” in response to Q13 (“What are your reasons for (currently) using AI technologies in your research?”; see Table 4). The qualitative analysis of the answers to the open follow-up question (Q14) shows similar thematic patterns seen in the answers to Q13. Participants say that they avail themselves of AI’s capacity to automate repetitive tasks such as coding, proofreading, translating, and pattern detection to save time, but normally verify outputs. A second reason concerns curiosity, as illustrated in this remark: “I learned how to use AI to know how technology changes (I am too young to start condemning change^^) and to teach my students how to use it in a sensible and critical way” (Q14/#38). A final, though less frequent, theme revolves around AI as a prospective research object in its own right as captured in this comment: “It’s not just a buddy/helper for research processes; it is a new research terrain in itself; it allows sociolinguistics to critically engage with new forms of datasets that AI produces (including biased, unethical etc. types of data)” (Q14/#7).

Participants also voiced much broader concerns about AI technologies in response to Q11, which asked why sociolinguists do *not* use AI (Q11). Table 5 shows that participants’ main reasons for not using AI tools such as LLMs are related to the perceived unreliability (56 %) and biased nature (55 %) of their outputs and data security (54 %) issues. Concerns over the representativeness of Big Data (24 %) and not seeing their usefulness (23 %) were also quite prominent reasons, in addition to other reasons, for not having engaged with them, such as a lack of time (20 %), expertise (13 %), or training (11 %). None of the participants said that they were not allowed by their institutions to use AI tools.⁵ In summary, this suggests that broader concerns about AI intersect with discipline-related motivations to constrain AI technology use.

Table 4: Percentage of respondents who (currently) use AI technologies in research for particular reasons.

Reason for using AI in research	%
Speed up time-consuming activities and increase efficiency	56 %
Use AI technologies out of curiosity	46 %
None of the above	23 %
Find the research support provided by AI very inspirational	16 %
Colleagues and collaborators encourage use	12 %
Detect new patterns that human analysis cannot identify	7 %
To analyse very large datasets	5 %
Told/counselled by employer to use AI	3 %
Because of the representativeness that the analysis of large datasets promises	1 %

⁵ The option “I have no access to AI tools” was not provided.

Table 5: Percentage of respondents who (currently) do not use AI technologies in research for particular reasons.

Reason for not using AI in research	%
The output is unreliable due to hallucination	56 %
The output is biased due to biased datasets	55 %
Concerns relating to data security	54 %
Don't think that Big Data represents human language	24 %
No need for using AI technologies in my work	23 %
Thinking about using AI but have not yet had time to engage with it	20 %
None of the above	13 %
Not sure how AI works and have no expertise in using it	13 %
Thinking about using AI but have not yet had a proper introduction	11 %
I don't know/I am not sure	1 %

These concerns were also confirmed by the 83 comments to the follow-up question (Q12 – “Please provide any additional information and comments relating to question 11”), where comments about the lack of reliability (38 %) and usefulness (23 %), and concerns of ethics and trust (39 %) were prevalent. Several respondents provided more than one reason for avoiding AI tools. Comment Q12/#71 states concerns over AI’s negative effect on the quality and essence of academic work in the humanities, including their own understanding of what being an academic entails, and the AI industry’s negative impact on the environment:

I don't use AI for writing, revising or editing because, for me, the joy of being an academic is working out what I want to say and how I want to say it. Farming those tasks out to a machine would be a huge diminution in my quality of (intellectual) life. [...] I also can't see myself jumping on the AI/LLM bandwagons in the near future because I am deeply worried about their cost to the planet in terms of their consumption of resources that humans need and their contribution to global warming. (Q12/#71)

Other comments express unease about “who owns this technology and, therefore, reaps the benefits (economic, political, etc.) from its adoption” (Q12/#10), that is, concerns about the commercial and power dimensions of AI control by large technology companies. As one respondent notes, “technologies are inherently unethical due to how data is gathered, how we're giving too much power to big tech” (Q12/#18). These reflections point to broader concerns about the concentration of technological ownership and the unequal distribution of benefits that arise from the adoption of AI.

Other ethical concerns include the usefulness of AI and how it may undermine important (academic) skills. One respondent wrote:

The data used to feed many large language models was collected unethically. I think using AI is actively harmful for one's critical thinking and nothing I've seen has convinced me that it's more useful than simply learning to do that function yourself. I also have serious environmental concerns. So I think it's just inherently unethical and negative. (Q12/#26)

Interestingly, non-use or a restricted use of AI is often not simply the result of conscientious objection but, prominently, disappointment with its capabilities and outputs, as discussed in the following comment:

I have used Chat GPT to generate summaries of others' research in fields that touch on my own. Some of these have provided useful starters. Others have been too superficial or “biased” (in the sense that Chat GPT made incorrect assumptions for the reasons for the request and produced summaries that would support those assumptions). I therefore remain somewhat sceptical about its value for me as a researcher – and therefore, for similar reasons, for my students. (Q12/#38)

While only a few respondents said that they did not know enough about the workings of AI to trust its output (Q12/#36), for others, it was knowledge about how AI works that made them not want to use it in their research:

I would not use it for analysis. AI tools/LLMs are about language prediction, they don't really “know” anything. They are prone to hallucinations. If I used AI tools to do analysis (e.g., code data), I would want to go back and check that it's all done correctly. And at that point, I might as well have just done it myself so it's not saving [...]. (Q12/#11)

Table 6: Percentage of respondents who report ethical and integrity issues associated with AI use.

Response to any perceived ethical or integrity issues around AI	%
Due to ethical or integrity issues, I use AI tools for specific purposes only	48 %
Due to ethical or integrity issues, I use AI tools as little as possible	29 %
Ethical or integrity concerns deter me from using AI	28 %
Ethical or integrity issues exist but do not hinder me from using AI tools	11 %
I am not aware of any issues serious enough to stop using AI tools	5 %

This quite dominant expression of doubt concerning the capacities of AI tools is further confirmed in responses to Q22, where we explicitly asked the extent to which ethical and integrity issues deterred respondents from using AI technologies. Table 6 shows that ethical awareness is an important factor that shapes sociolinguists' engagement with AI. Nearly half of respondents (48 %) report being aware of such issues and, as a result, limited their use of AI tools to specific supportive purposes. A further third use AI as little as possible due to these concerns, and another third state that ethical or integrity considerations prevent them from using AI altogether. Only a small percentage (11 %) acknowledge these issues but claim they do not restrict their AI use. An even smaller proportion (5 %) share no awareness of any serious ethical risks. The interest in this issue is confirmed by the fact that we received 137 often rather lengthy answers – the longest consisting of 424 words – to an open-ended question (Q21) that asked participants to reflect on ethical issues.⁶ In sum, most respondents exercise caution and restraint when engaging with AI tools due to ethical and integrity considerations; this may partly explain the limited adoption of AI technologies in current sociolinguistics research.

4 Conclusions

In the context of fierce competitiveness driven by the neoliberal university, a sector seemingly condemned to the consequences of chronic underfunding, researchers are faced with the dilemma of using AI tools amid big tech's inflated marketing narratives of disruption, institutional expectations, personal moral imperatives, and disciplinary epistemological allegiances. Under structural pressures to maximize productivity and efficiency, sociolinguistics emerges as a prudent field, calling for critical pondering on the so-called AI hype.

Use of AI has become popular but is met with considerable scepticism, especially regarding its application in analysis and research. According to our data, many core sociolinguists make some use of AI to deal with time-consuming and repetitive tasks when preparing their research and in their wider academic work, confirming a broader societal trend. Yet perhaps the more significant finding is that this pragmatically oriented engagement is met with vocal calls for caution, if not outright resistance, from many mainstream sociolinguists who raise concerns about ethics, the environment, research integrity, and power dynamics. Referring to the theme of this special issue, our data do not indicate that AI technology has reached a point of disruption in (mainstream) sociolinguistics research.

Another noteworthy aspect is the significant number of participants who fully engaged with the survey and, more specifically, the unusually high response rate to the open-ended questions and the richness of the responses. This is significant in itself and may partly be due to our methodological choices, including the many follow-up prompts inviting elaboration on closed questions. However, we believe this engagement provides strong evidence of the scholarly interest and intellectual investment that AI currently sparks among preoccupied sociolinguists. This is probably unsurprising given the high level of sociopolitical uncertainty and sociolinguistics' deep commitment to principles of equity and social justice.

⁶ "Please tell us here if you see any ethical or integrity concerns in using AI for research and wider academic activities, particularly regarding issues such as privacy, environmental impact and linguistic representativeness (both within and across languages)?".

In sum, the increasing – but still limited – technologization of research does not appear to have altered prevailing understandings of what constitutes sociolinguistics research. A promising future avenue of inquiry is to explore how AI may be shaping our current understandings of language in society and their epistemological implications.

References

- Androutsopoulos, Jannis. 2016. Theorizing media, mediation and mediatization. In Nikolas Coupland (ed.), *Sociolinguistics: Theoretical debates*, 282–302. Cambridge: Cambridge University Press.
- Erdocia, Iker, Bettina Migge & Britta Schneider. 2024. Language is not a data set – Why overcoming ideologies of dataism is more important than ever in the age of AI. *Journal of Sociolinguistics* 28(5). 20–25.
- Erdocia, Iker, Britta Schneider & Bettina Migge. 2026. Language in the age of AI technology: From human to non-human authenticity, from public governance to privatised assemblages. *Language in Society* 55(3). 523–543.
- Grieve, Jack, Sara Bartl, Matteo Fuoli, Jason Grafmiller, Weihang Huang, Alejandro Jawerbaum, Akira Murakami, Marcus Perlman, Dana Roemling & Bodo Winter. 2025. The sociolinguistic foundations of language modeling. *Frontiers in Artificial Intelligence* 7. <https://doi.org/10.3389/frai.2024.1472411>.
- Hao, Karen. 2025. *Empire of AI: Dreams and nightmares in Sam Altman's OpenAI*. New York: Penguin Press.
- Kelly-Holmes, Helen. 2022. Sociolinguistics in an increasingly technologized reality. *Sociolinguistica* 36(1–2). 99–110.
- Kelly-Holmes, Helen. 2024. Artificial intelligence and the future of our sociolinguistic work. *Journal of Sociolinguistics* 28(5). 3–10.
- Koenecke, Allison, Andrew Namb, Emily Lake, Joe Nudell, Minnie Quartey, Zion Mengesha, Connor Toups, John R. Rickford, Dan Jurafsky & Sharad Goel. 2020. Racial disparities in automated speech recognition. *Proceedings of the National Academy of Sciences* 117, 7684–7689.
- Lamoureaux, Siri, Michael Castelle & Anna Weichselbraun. 2026. Language machines: Toward a linguistic anthropology of large language models. *Journal of Linguistic Anthropology* 36(1). <https://doi.org/10.1111/jola.70033>.
- Maly, Ico. 2022. Algorithms, interaction and power: A research agenda for digital discourse analysis. *Working Papers in Urban Language & Literacies* 295. 2–26.
- Maly, Ico. 2024. AI, power and sociolinguistics. *Journal of Sociolinguistics* 28. <https://doi.org/10.1111/josl.12681>.
- Markl, Nina. 2022. Language variation and algorithmic bias: Understanding algorithmic bias in British English automatic speech recognition. In *FACCT '22: Proceedings of the 2022 ACM conference on fairness, accountability, and transparency*, 521–534. New York, NY: Association for Computing Machinery.
- Meyerhoff, Miriam. 2011. *Introducing sociolinguistics*. London: Routledge.
- Nguyen, Dong. 2025. Collaborative growth: When large language models meet sociolinguistics. *Language and Linguistics Compass* 19. <https://doi.org/10.1111/lnc3.70010>.
- Nguyen, Dong, Seza A. Dogruöz, Carolyn P. Rosé & Franciska de Jong. 2016. Computational sociolinguistics: A survey. *Computational Linguistics* 42(3). 537–593.
- Paulus, Trena, Jessica Nina Lester & Charlie Davis. 2026. The construction of the role of AI in qualitative data analysis in the social sciences. *AI & Society* 41. 1737–1748.
- Perez-de-Arriluzea-Madariaga, Amaia. 2025. Computational sociolinguistics: An emerging multidisciplinary research area. *International Journal of Applied Linguistics* 35(3). 1298–1313.
- Savoldi, Beatrice, Marco Gaido, Luisa Bentivogli, Matteo Negri & Marco Turchi. 2021. Gender bias in machine translation. *Transactions of the Association for Computational Linguistics* 9. 845–874.