The Global Digitalization of the University and Language in the Era of APT: A Prompt Response to "AI and the University as a Service"

Lee, Ji Eun [*]

0000-0002-4054-8981

Abstract

This essay is a prompt response to Matthew Kirschenbaum and Rita Raley's essay "AI and the University as a Service" published in the most recent issue of *PMLA*. I diagnose the anticipated problem of higher education in the era of AI as the "global digitalization of the university and language." By the digitalization of the university, I mean the online circulation of recorded courses. By the digitalization of language, my approach broadens to refer not only to the proliferation of online resources but also the dissemination of language into machine-readable units and their redistribution. My positionality in interpreting these kinds of digitalization is global, involving multilingual, multicultural contexts. The possibilities and challenges concern the use of LLMs for language education and writing pedagogy in higher education in two aspects: 1) whether it is justifiable to train students using the language of LLM outputs based on statistical probability rather than communicative intent; 2) how to decolonize the English monopoly on language education and data structures. My suggestions are open-ended, reminding us humanists of the Korean resources that are widely accessible and the co-evolutionary approach of writing with AI. All we need in future higher education may be APT (AI Personal Training).

Keywords: AI, LLM, Language, Higher Education, Digitalization, APT (AI Personal Training)

1. Introduction

What if digitalization is not a specific condition required for digital humanities but "a general condition of language and life" (508) that will prevail in the future of humanities research and education? This is what Matthew Kirschenbaum and Rita Raley alert us to in their essay "AI and the University as a Service" published in the most recent issue of *PMLA*.[1] As they suggest, the computational approaches that characterized the journal's previous Theories and Methodologies section on digital humanities in January 2020 (135.1)[2] "are increasingly regarded as altogether ordinary" (508), while the imminent structural changes in the humanities with the advent of LLMs, such as ChatGPT in November 2022, look unprecedented.

I would like to respond to their call to rethink the future of humanities research and education in the era of large language models and generative AI from my position as an English PhD who studied writing (and thinking!) in English through both undergraduate and graduate education in Korea and the United States. As someone who first learned English in Korea and later in the United States, I will focus on the *digitalization* of the university and *language*, especially from a *global* perspective. By the digitalization of the university, I mean mostly the courses

[*] 1st Author: Department of English Language and Literature, Sungkyunkwan University, Sungkyunkwan-ro 25-2, Seoul 03063, jieunclee@skku.edu

recorded, uploaded, and shared on digital platforms such as Coursera, edX, Udacity, etc. In the case of the digitalization of language, the meaning of the word "digitalization" refers to not only the proliferation of online resources for language learning but also, and more importantly, to the reformulation of language into machine-readable units through tokenization, word embedding, etc., and the production of sentences by prediction. This is the language pervading LLMs such as ChatGPT, Gemini, and HyperCLOVA. My positionality in interpreting these kinds of digitalization is global, involving multilingual, multicultural contexts.

My focus is mostly on writing pedagogy, especially because what distinguishes language education in the university from that of the private sector is the academic goal shared by all disciplines in the humanities to develop the capacity for critical thinking and writing. Writing in one's native language requires a lot training, and L2 writing, or writing in a language other than one's native language, requires not only the acquisition of language skills but also the reformulation of the thinking process in terms of culture. With the digitalization of language, university education calls for new pedagogical adaptation. As Matthew Kirschenbaum and Rita Raley write, "Language—in the form of executable code, machine-readable data, and humanreadable text subject to tokenization for word embeddings—is not merely product but also managerial apparatus (yes, *dispositif*) that produces humans as subjects of, and subjects to, the extractive economies that are the new scene of writing" (512).

2. The Global Digitalization of the University and Language: Possibilities and Challenges

The pandemic and its enforcement of online lectures accelerated the digitalization of the university as a service, which is now at hand with the rise of shared online platforms that store and stream recorded lectures. Many universities in Korea as well as those abroad have created online platforms to share their lectures within their university communities and sometimes outside of them. Besides universal online platforms that stream and archive courses such as Coursera, edX, Udacity, Stanford Online, MIT OpenCourseWare (OCW), Harvard Online Learning, there are also Korean versions such as SNU OpenCourseWare (OCW), Korea University OpenCourseWare (KUOCW), SKKU Online, to name a few. While these online platforms transform the originally in-person, communicative, communal forms of the class setting into repetitive, receptive, and isolating forms, they also make higher education accessible beyond the barriers of physical distance. For my graduate seminar on academic writing in English last semester, for example, I used the recorded workshops from UCLA graduate writing center (https://gwc.gsrc.ucla.edu/workshop-videos) to teach about the academic norms of writing circulating in Anglophone academia in the form of flipped learning. Students watched the recorded workshops in advance of class and came to class with drafts to peer-review and revise. This kind of global digitalization of university lectures enhances international accessibility to higher education and makes the academia inclusive of diverse cultures and nations.

On the other hand, the digitalization of language at first glance seems to refer to the same trend of global accessibility to language resources that are abundant in vernacular environments available on the internet, especially with the growing use of LLMs for language learning in the 2020s. This changing accessibility to language learning, however, is underpinned by the digitalization of language in another sense: the computational processing of language into machine-readable units, known as natural language processing (NLP). NLP is a field of computer science focused on enabling computers to analyze and understand human language. It involves technologies such as text preprocessing (e.g. tokenization, cleaning, normalization, stemming, padding) and linguistic representations at lexical, syntactic, and semantic levels (e.g. part-of-speech tagging, topic modeling). These NLP methods are foundational to language models (LMs), which are systems trained to predict sequences of words. Language models have evolved from N-gram models–a sequence of n-words–and word-vectors, then to Recurrent Neural Networks (RNN), which were capable of using "memory" of previous inputs for sequence

prediction and text classification.[3] Most recently, Large Language Models (LLMs) of Transformer architecture like ChatGPT, released in November 2022, use self-attention mechanism to perform more complex tasks like summarization and question-answering.

While the digitalization of language seems helpful to students who learn foreign languages, especially English, across the globe, it poses some challenges. First, is it justifiable to train students on the language of LLM outputs based on statistical probability rather than communicative intent, which, according to Steven Knapp and Walter Benn Michaels in their essay "Against Theory,"[4] constitutes meaning in human language? Given that an LLM is "a stochastic parrot" producing words "without any reference to meaning" (617), it is questionable whether an L2 learner who studies the language solely figured as probable patterns could attain, or should attain, the same capacity of contextualizing knowledge and communicative interaction. As the authors of "On the Dangers of Stochastic Parrots" write, "languages are systems of signs, i.e. pairings of form and meaning. But the training data for LMs is only form; they do not have access to meaning" (615). LM outputs, which consist of "linguistic forms" produced by statistical probability without any reference to meaning, may trick us into believing that they constitute real language and demand human understanding. Bender et al. caution that an LM is "a stochastic parrot" that spreads out a string of words without meanings, and humans who try to read and interpret it may be "led down to the garden path" (617, 616). A more serious concern is that models only trained on LLM outputs could collapse, suggesting that machine-produced texts are not good enough to be the source of language learning in the case of language models themselves,[5] which may or may not apply to human brains as well. It has been discovered that machine-produced texts show different patterns of probability among word-distributions.[6] Will humans' linguistic capacity trained on inputs produced by LLMs collapse as well?

Secondly, the vast size of English datasets and the limited access to texts in other languages in most big-tech LLMs may reinforce Anglocentrism and the English monopoly on datasets rather than promote multilingualism in higher education. I second Eduardo Ledesma's argument that AI and LLMs could exacerbate "the existing marginalization of languages 'other' than English, by buttressing English dominance and further eroding linguistic diversity in higher education and beyond" (534).[7] Although his review of the effects of English monolingualism through AI and LLMs mostly relates to the Global South, this issue is not significantly different in East Asia and for Asian languages. When the "HyperCLOVA X Technical Report" was first released by NAVER in April 2024,[8] people expected it to have good multilingual skills as said in the report,[9] but many found HyperCLOVA X's Korean performance very disappointing. When I used it, it produced sentences that stop with subwords due to the word limit, and its style, tone, and diction were too casual. When I asked Clova how it learned Korean, it said it only read casual or informative writings such as blog posts or news articles, not the academic articles available in DBpia. The limitations of Korean LLMs were already confirmed by Joo et al. in 2023.[10] When LLaMA and Polyglot become available as open source, Korean models were developed using pre-trained LLMs as foundation models: Polyglot-ko modeled on Polyglot, KoAlpaca-Polyglot and KULLM modeled on Polyglot-ko and KoAlpaca modeled on LLaMa, and KoVicuna modeled on Vicuna.[11] However, the size of their Korean parameters as well as that of HyperCLOVA X, however, is far smaller than the size of English parameters available for GPT 4o.[12] Users of GPT may all have experienced that its answers about Korean culture and history are often inaccurate or fabricated. To resist this "new form of coloniality" (which Ledesma uses to refer to this problem of English dominance over other languages 354), it is important for each nation to develop a vernacular language model that is attuned to its vernacular culture and history, but the kind and size of current datasets used by Korean LLMs for these decolonizing tasks are far from enough.

3. Moving Forward Writing with AI, or APT (AI Personal Training)

What should we do to move forward in the face of this global digitalization of the university and language driven by AI? Echoing Katherine Elkins, I would say that the question is not how or whether we "serve" this new apparatus called AI that restructures higher education but about how to "advocat[e] for an AI that would serve us" (546).[13] Responding authors to Kirschenbaum and Raley's thought-provoking essay reflect on the question of language and offer various approaches and suggestions for "AI and the University as a Service." As Aarthi Vadde tells us, human-centered writing is no longer the default when more and more students are tempted to use generative AI to complete their writing assignments and as automatic text-generation proliferate in daily life. Vadde adds that we humanists, as called upon to help establish the university's task force and guidelines for writing, are reminded that "our domain is language" (557). Meredith Marin presents "Command Lines for the Humanities," calling for collaboration, communication, method, and collective approaches to datasets and data reviewing.[14] Seth Perlow urges us to restore "aesthetic criticism" to define the contours of "value." Similarly, Timothy Laquintano and Annette Vee highlight the potential "linguistic imperialism" and the validity of the "human voice as a marker of value," inviting us to reflect on the history of discrimination and hierarchies to better reshape "writing voice" of or between humans and machines. Their emphasis on the value-judgements that define what should be counted as "good" output, good writing, good analysis is absolutely needed in the contemporary textual vista where "textpocalypse" is near.[15] As I mentioned earlier, Eduardo Ledesma invites non-English studies scholars to engage in shaping critical AI studies from diverse perspectives in order to decolonize the Anglocentrism that pervades not only LLMs and literary studies but also other disciplines, implying that AI may be "another cultural form altogether" (538). Junting Huang discusses writing as a "therapeutic practice" drawing upon Wittgenstein's idea of language as an active, practical utterance in the time of crisis.[16] Holly Dugan and Dolsy Smith refuses to debate which reading is better between humanist reading and computational reading and instead welcomes "the cultivation of the lag that this pedagogy introduces and the critique that it hopefully inspires in its wake: a hiatus or suspension" (521).

Hoping on this train of thoughts and suggestions for AI's role in higher education, I will suggest what I think about the challenges posed by the global digitalization of the university and language, which I have identified as the problem of the future higher education anticipated in Kirschenbaum and Raley's caricature. The second question about coloniality requires an ongoing reformulation of policies and cooperation between the national governments and LLM developers. In the United States, many current literary archives storing texts protected by intellectual property are closed to public access in the name of fair use, as seen in Warhol v. Goldsmith, as Seth Perlow laments in his response essay because enough training in large data is necessary for the consecutive development of AI models.[17] Access to library resources and news archives in Korea, however, is relatively flexible and open because most textual data in Korea are funded by the government and are open to the public or members of academic institutions for fair use. The Korean Association for Digital Humanities (KADH) has a webpage for data available in Korean or other languages: https://www.kadh.org/category/resources/data/. Besides those tailored datasets, current Korean textual data available in ancient and contemporary Korean include the following:[18]

Table 1. Korea Textual Data Archives

Krpia(Korean DATABASE, 한국의 지식콘텐츠) 중 역사/ 철학
(https://www.krpia.co.kr/subjectclassify/krpiaList?classifyName=krpia&categoryCode=4)
한국고전종합DB(https://db.itkc.or.kr/),
고전용어 시소러스(http://thesaurus.itkc.or.kr/)
조선왕조실록 DB(https://sillok.history.go.kr/main/main.do)

국사편찬위원회 한국사데이터베이스 (https://db.history.go.kr/)

한국학중앙연구원 한국학통합자료플랫폼 (https://kdp.aks.ac.kr/)

국립중앙도서관 국가자료종합목록 (https://www.nl.go.kr/kolisnet/index.do)

국립중앙도서관 대한민국신문아카이브 (https://nl.go.kr/newspaper/)

국립국어원 모두의말뭉치 (https://kli.korean.go.kr/corpus/main/requestMain.do?lang=ko)

국사편찬위원회 3.1운동 데이터베이스 (https://db.history.go.kr/samil/)

데이터로 다시 읽는 조선시대 양반의 생활(지암일기 DB) (https://jiamdiary.info/)

공공데이터포털 (https://data.go.kr): 분류체계, 서비스유형, 확장자로 필터링하여 문헌자료 검색 가능. 그 예로 **한국고전번역원 고전원문**, **한국학중앙연구원 장서각원문이미지 분류정보통합관리**가 있음.

기억의 도서관 코리안 메모리 (https://koreanmemory.nl.go.kr)

Of course, not all Korean data are available or free of copyright laws, and I am just pointing out the relative generosity of public accessibility in South Korea compared to the aggressive copyright laws in the States. Korean LLMs such as HyperCLOVA X should leverage this large amount of original Korean textual data widely accessible and available as well as academic language data available in DBpia (https://www.dbpia.co.kr/), just as GPT uses ScholarAI and other plug-in services (though some of them are to be paid additionally) for that purpose. Another reason that Korean LLMs are less developed than English LLMs may probably be the more complex morphology required for tokenization, including the use of particles (josa) that classifies word phrases into grammatical categories such as noun phrases, verb phrases, etc. More complex NLP tools such as KIWI (Korean Intelligent Word Identifier) should be used. Min-Chul Lee's article "Kiwi: Developing a Korean Morphological Analyzer Based on Statistical Language Models and Skip-Bigram," which examines how to resolve ambiguities in morphological analysis, has recorded more than 700 views, showing high interest in Korean linguistics and textual data.[19]

As for the first challenge about whether it is justifiable to train students using the language of LLM outputs based on statistical probability rather than communicative intent, I would argue we humans are born to read and the question is not about justification but about how to develop new *hermeneutics* that will help us navigate this new landscape of reading and writing called "textpocalyse," where texts produced by humans and/or machines without intent proliferate and blend together.[20] Citing Kittler's words on "*hermeneutics*–a pilot's understanding of signals, whether they stem from gods, machines, or sources of noise," Avery Slater argues that it is necessary to develop the capacity to interpret, or to make sense of, multiple forms of signals that are not explicitly generated by human intention (quoted from Kittler. "Signal-to-Noise Ratio" 177).[21] We humans "read [. . .] literally" anyway when we see a text, but this reading is different from "interpretation"—an activity that gauges the meaning of the text.[22] We should distinguish between mechanical reading that machines can do as well and subjective interpretation that only belongs to humans, as Tyler Shoemaker alerts us to in his essay "Machines, Reading." All in all, in the age of textpocalypse and with LLMs being commercialized for language learning business, humans will read anyway and they will keep interpreting whether or not the text they see is produced by humans with intentions or machines without intentions. By reading it and interpreting it anyhow, we humans can transform the string of linguistic forms absent of reference into the language pairing forms and meanings. This act of interpretation out of nothing, however, was also what humanists have done for decades, as post-structuralists like Roland Barthes free interpretation from the author's intention in the "Death of the Author."[23] Reading and interpreting machine-written texts as learning materials would work as long as humans are aware of it and know how to identify "disinformation" (Slater 1258), because, unlike machines that process linguistic patterns without a reference, humans can dissect the language and re-pair its forms with meanings.[24]

Equipped with the capacity of interpretation, we humanist scholars and students can read and write *with* AI, as if we have AI Personal Training (APT) that helps us with writing, rather than either totally disregarding or relying

on it. Much has been said about how to read with AI by scholars who developed diverse methods of distant reading, and their conclusion is to keep both close-reading and distant-reading as main methods and not to lose our humanistic perspectives that are attuned to localized contexts. But not much has been said about how to *write with* AI in academia yet, especially in L2 context. Though not many academic articles are available on this topic, some have reviewed and recommended ways of collaborating with AI in academic writing.[25] In Korea, this new task of how to write with AI, especially with LLMs on the level of higher education, is more actively discussed in publishing industries that promote practical approaches. Tae-Yong Kim provides a variety of approaches and prompts that could be used to proofread and correct grammatical errors.[26] Last semester when I taught a graduate writing seminar, for example, I let my students use ChatGPT to correct their grammar, and their learning was more effective when their English was less fluent. GPT was mostly correct in its revisions except in a few cases when a nuanced understanding of the context was required, which attests to the fact that it still requires human attention and proofreading. My students confirmed that GPT 3.5, which was free, judged ungrammatical sentences as correct and did not correct anything, while GPT 4.0, which required subscription fees, made multiple corrections on the same text.

The partnership with AI and LLMs to figure out probable distributions of words for precise diction predates the advent of GPT. Language learners in academic settings and university writing centers have used large corpora like the Corpus of Contemporary American English (COCA), the British National Corpus (BNC), or WordNeighbors and electronic software by checking n-strings of words and their probable collocations. Mobile-Assisted Language Learning (MALL), which refers to the type of language learning that uses laptops, tablets, and phones to access online resources, was initially recognized in the late 1990s and has been actively discussed in contemporary education studies. ESL students in American universities showed great development in L2 writing sills by using grammatical applications, especially in their use of verbs and prepositions.[27] Eun-Young Kwon's study of academic trends in MALL studies in KCI articles, which used LDA topic modeling and word cloud to read 247 LACI articles on MALL, show that the word "university" or "college" appears significantly more (eighty-six, sixty-eight times respectively) than words of secondary or elementary or pre-elementary schools (twenty-one times respectively), indicating that MALL is more used in higher education.[28] Taecharungroj (2023) and Sallam (2023) even encourage the use of GPT for the whole process of writing including brain storming, researching, editing, and proofreading, arguing that it is tremendously useful. Yet a more nuanced sense of academic integrity[29] is absolutely required, if the use of LLMs exceeds beyond subtle grammatical corrections.

In this context, I propose APT, by which I mean "AI Personal Training," a practice involving humans partnering with AI to train, or improve, their writing.[30] My suggestions regarding APT include, first, the practical realm of developing writing pedagogy, which encompasses multiple steps of thinking or brainstorming and improving L2 writing skills, as outlined above. People now find LLMs such as GPT and Gemini very helpful for brainstorming ideas, finding research topics, and conducting literature reviews, though the summaries they provide are not always accurate. While it may be acceptable to get help from APT for brainstorming and editing, however, I believe that drafting should solely be done by humans, relying on their own capacities.

Secondly, and more importantly, while it is crucial to maintain human agency in writing, in the era of APT, we humans should embrace co-evolution with AI in the realm of language. When the languages humans read in contemporary environments are already influenced by chatbots, machine-translated subtitles on Zoom, and generative text models, human and machine languages will evolve by influencing one another. The co-evolution between different species, which Donna Haraway has introduced when the binary between nature and culture no longer holds in the co-habitable space shared by humans and dogs,[31] is applicable to humans' relationships with AI. If we humans dare to believe in our hermeneutic capacity, we may venture into this new arena where human and machine-generated languages co-evolve with each other, shaping our co-constitutive relationships and languages.

4. Coda

The meaning of the word "Prompt" in the subtitle of my essay is twofold. Firstly, it is an ambiguously belated yet prompt response to Kirschenbaum and Raley's essay, as my ideas for this essay were brewing but were never put into words because of my busy schedule when I was asked to write a quick response to their essay in early March with a very short deadline. I hope my response, though belated, is prompt enough to be printed in the same year as their essay's publication. Secondly, I hope my essay becomes another prompt for future responses from the non-Anglophone worlds, reflecting on the possibilities and challenges that the LLMs and AI bring to other languages and universities spoken, written, or located outside Anglophone academia. Unlike Kirschenbaum and and Raley's concluding remarks—"AI is not the prompt, it is the punctuation"—I would say, "AI is the prompt," especially for APT (AI Personal Training), which will dominate or become common in future higher education.[32] Perhaps, what Kazuo Ishiguro described in his 2021 novel *Klara and the Sun* is becoming or has already become real. Don't we all have an "Artificial Friend" like Klara, though in the form of a disembodied voice like ChatGPT, installed on our cellphones? Alas, not everyone but only those who can afford to pay the subscription will have APT.

---

[1] Kirschenbaum, Matthew, and Rita Raley. "AI and the University as a Service." *PMLA/Publications of the Modern Language Association of America* 139, no. 3 (2024): 504–15. https://doi.org/10.1632/S003081292400052X.

[2] *PMLA*, "Special Topic: Varieties of Digital Humanities," vol. 135, no. 1 (January 2020). https://www.cambridge.org/core/journals/pmla/issue/847B9F9004B8462A86F0323153F760B1.

[3] For this definition of natural language processing and its significance to language models, I referred to the following: Jacob Eisenstein, *Introduction to Natural Language Processing* (Cambridge, MA: MIT Press, 2019); Wonjun Yoo and Jun Sang, "Introduction to Natural Language Processing Using Deep Learning", 2024, accessed December 13, 2024, https://wikidocs.net/book/2155; YunSeok Choi, "Introduction to Natural Language Processing" (undergraduate course, Sungkyunkwan University, South Korea, Fall 2024); and Emily M. Bender, Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell, "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?" in *FAccT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (New York: Association for Computing Machinery, 2021), 610–23, https://dl.acm.org/doi/10.1145/3442188.3445922.

[4] Steven Knapp and Walter Benn Michaels, "Against Theory," *Critical Inquiry* 8, no. 4 (1982): 723–742, https://doi.org/10.1086/448178.

[5] Adina Bresge, "Training AI on Machine-Generated Text Could Lead to 'Model Collapse,' Researchers Warn," *University of Toronto News*, August 21, 2023, https://www.utoronto.ca/news/training-ai-machine-generated-text-could-lead-model-collapse-researchers-warn.

[6] Alberto Muñoz-Ortiz, Carlos Gómez-Rodríguez, and David Vilares, "Contrasting Linguistic Patterns in Human and LLM-Generated News Text," *Artificial Intelligence Review* 57 (2024): 265–292, https://doi.org/10.1007/s10462-024-10903-2. The difference between human writings and AI-generated texts are summarized in the abstract as follows: "Human texts exhibit more scattered sentence length distributions, more variety of vocabulary, a distinct use of dependency and constituent types, shorter constituents, and more optimized dependency distances. Humans tend to exhibit stronger negative emotions (such as fear and disgust) and less joy compared to text generated by LLMs, with the toxicity of these models increasing as their size grows. LLM outputs use more numbers, symbols and auxiliaries (suggesting objective language) than human texts, as well as more pronouns. The sexist bias prevalent in human text is also expressed by LLMs, and even magnified in all of them but one. Differences between LLMs and humans are larger than between LLMs."

[7] Seth Perlow. "Generative Theories, Pretrained Responses: Large AI Models and the Humanities." *PMLA/Publications of the Modern Language Association of America* 139, no. 3 (2024): 548–52. https://doi.org/10.1632/S0030812924000518.

[8] NAVER Cloud, HyperCLOVA X Team, "HyperCLOVA X Technical Report," *arXiv*, April 2024, https://arxiv.org/abs/2404.01954.

[9] Maeil Business Newspaper, "Announcement of the Technical Report emphasis on multilingual skills 'Prove global competitiveness,'" *Maeil Business Newspaper*, April 4, 2024, https://www.mk.co.kr/en/it/10982237

[10] 주하영, 오현택, 양진홍. (2023). 오픈 소스 기반의 거대 언어 모델 연구 동향: 서베이. 한국정보전자통신기술학회 논문지, 16(4), 193-202. Ha-Young Joo, Hyeontaek Oh, & Jinhong Yang (2023). A Survey on Open Source based Large Language Models. Journal of Korea Institute of Information, Electronics, and Communication Technology, 16(4), 193-202. https://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE11506783 Also see Haziqa Sajid, "The State of Multilingual LLMs: Moving Beyond English," *Unite.AI*, February 10, 2024, https://www.unite.ai/the-state-of-multilingual-

llms-moving-beyond-english/.

[11] See Table 3 in Joo et al. Most of the Korean parameters used by Korean LLMs mentioned here are translated from English. KMMLU is said to be the first MMLU that consists of original Korean texts (not the texts translated from English). HyperCLOVA X showed promise in its use of original Korean texts, but it cannot access DBpia.

[12] Josh Howarth, "Number of Parameters in GPT-4 (Latest Data)," *Exploding Topics*, August 6, 2024, https://explodingtopics.com/blog/gpt-parameters.

[13] Katherine Elikins. "A(I) University in Ruins: What Remains in a World with Large Language Models?" *PMLA/Publications of the Modern Language Association of America* 139, no. 3 (2024): 559–65. https://doi.org/10.1632/S0030812924000543.

[14] Meredith Martin. "Command Lines for the Humanities." *PMLA/Publications of the Modern Language Association of America* 139, no. 3 (2024): 541–47. https://doi.org/10.1632/S0030812924000555.

[15] Matthew Kirschenbaum, "Prepare for the Textpocalypse," *The Atlantic*, March 8, 2023, https://www.theatlantic.com/technology/archive/2023/03/ai-chatgpt-writing-language-models/673318/.

[16] Huang, Junting. "Large Language Games, Therapeutic or Otherwise." *PMLA/Publications of the Modern Language Association of America* 139, no. 3 (2024): 522–26. https://doi.org/10.1632/S0030812924000531.

[17] Perlow, ibid. 550-52.

[18] Many thanks to Jaehyun Seo and Joochan Choi for helping me to make this list

[19] Min-Chul Lee, "Kiwi: Developing a Korean Morphological Analyzer Based on Statistical Language Models and Skip-Bigram," *Korean Journal of Digital Humanities* 1, no. 1 (2023): 109–135, https://accesson.kr/kjdh/v.1/1/109/43508.

[20] Matthew Kirschenbaum, "Prepare for the Textpocalypse," *The Atlantic*, March 8, 2023, https://www.theatlantic.com/technology/archive/2023/03/ai-chatgpt-writing-language-models/673318/

[21] Avery Slater, "Hermenautics: Toward a Disinformation Theory," *New Literary History* 54, no. 2 (2023): 1258, https://doi.org/10.1353/nlh.2023.a907171.

[22] Tyler Shoemaker, "Machines, Reading," *Critical Inquiry* (blog), June 28, 2023, https://critinq.wordpress.com/2023/06/28/machines-reading/

[23] See Ted Underwood in his review of "On the Dangers of Stochastic Parrots" and Knapp and Michaels's "Against Theory," and Seth Perlow's summary of Underwood's arguments in his essay (550)

[24] Of course, my prediction should be confirmed by empirical experiences in the long run, but for now, I believe in the humanist practice of interpretation

[25] See, for example, Zadorozhnyy and Lai (2023), Jacob (2024), and Soo-Hyung Joo (2024).

[26] Tae-Yong Kim, *Research Methods in Social Sciences Using Artificial Intelligence* (Seoul: Kyobo Book Centre, 2023), https://product.kyobobook.co.kr/detail/S000213184420. While I do not agree with everything he suggests about how to use AI for academic writing, I think the use of AI for proofreading works well, with cautions.

[27] Li and Hegelheimer, 2013.

[28] Eun-Young Kwon. 2024. "Analysis of trends in domestic mobile-assisted language learning in English education: Using text mining." Korea Journal of English Language and Linguistics 24: 325-347. 10.15738/kjell.24..202404.325

[29] Many universities are working to provide reliable guidelines for using generative AI in academic writing without losing academic integrity. Some efforts have been made in universities in Korea as well as those abroad. See 성균관대학교 챗GPT 종합안내 홈페이지; 고려대학교 ChatGPT 활용 가이드라인; and for more examples in Korea, check https://aiedap.or.kr/?page_id=112&mod=document&uid=338. Some examples from universities abroad include the following: https://academicintegrity.ubc.ca/chatgpt-faq; https://oai.missouri.edu/chatgpt-artificial-intelligence-and-academic-integrity/

[30] There are many AI personal training tools that assist humans with health and workout plans; however, my use of APT refers to assistance with writing.

[31] Donna Haraway, *The Companion Species Manifesto: Dogs, People, and Significant Otherness* (Chicago: Prickly Paradigm Press, 2003)

[32] This echoes Rosé and Bruno Mars' recent song "APT," which I think reverses the universal Anglophonism by confidently shouting the word in Konglish. The song itself deserves a whole new essay.

References

Bender, Emily M., Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?" In *FAccT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–23. New York: Association for Computing Machinery, 2021. https://dl.acm.org/doi/10.1145/3442188.3445922.

Review

Bresge, Adina. "Training AI on Machine-Generated Text Could Lead to 'Model Collapse,' Researchers Warn." *University of Toronto News*, August 21, 2023. https://www.utoronto.ca/news/training-ai-machine-generated-text-could-lead-model-collapse-researchers-warn.

Choi, YunSeok. "Introduction to Natural Language Processing." Undergraduate course, Sungkyunkwan University, South Korea, Fall 2024.

Dugan, Holly, and Dolsy Smith. "Untimely Models." *PMLA/Publications of the Modern Language Association of America* 139, no. 3 (2024): 516–21. https://doi.org/10.1632/S0030812924000592.

Elkins, Katherine. "A(I) University in Ruins: What Remains in a World with Large Language Models?" *PMLA/Publications of the Modern Language Association of America* 139, no. 3 (2024): 559–65. https://doi.org/10.1632/S0030812924000543.

Geher, Glenn. "ChatGPT, Artificial Intelligence, and the Future of Writing." *Psychology Today*, January 2023. Accessed November 20, 2023. https://www.psychologytoday.com/us/blog/darwins-subterranean-world/202301/chatgpt-artificial-intelligence-and-the-future-of-writing.

Haraway, Donna. *The Companion Species Manifesto: Dogs, People, and Significant Otherness*. Chicago: Prickly Paradigm Press, 2003.

Howarth, Josh. "Number of Parameters in GPT-4 (Latest Data)." *Exploding Topics*, August 6, 2024. https://explodingtopics.com/blog/gpt-parameters.

Huang, Junting. "Large Language Games, Therapeutic or Otherwise." *PMLA/Publications of the Modern Language Association of America* 139, no. 3 (2024): 522–26. https://doi.org/10.1632/S0030812924000531.

Jacob, Sharin R., Tate, Tamara and Warschauer, Mark. "Emergent AI-assisted discourse: a case study of a second language writer authoring with ChatGPT" *Journal of China Computer-Assisted Language Learning* (2024). https://doi.org/10.1515/jccall-2024-0011

Ha-Young Joo, Hyeontaek Oh, & Jinhong Yang (2023). A Survey on Open Source based Large Language Models. Journal of Korea Institute of Information, Electronics, and Communication Technology, 16(4), 193-202. 주하영, 오현택, 양진홍. (2023). 오픈 소스 기반의 거대 언어 모델 연구 동향: 서베이. 한국정보전자통신기술학회 논문지, 16(4), 193-202.

Joo, Soo Hyoung. "Generative AI as Writing or Speaking Partners in L2 Learning: Implications for Learning-Oriented Assessments." *Studies in Applied Linguistics & TESOL at Teachers College, Columbia University* 24, no. 1 (2024): 54–59. https://doi.org/10.1177/00336882231224813.

Kim, Tae-Yong. *Research Methods in Social Sciences Using Artificial Intelligence*. Seoul: Kyobo Book Centre, 2023. https://product.kyobobook.co.kr/detail/S000213184420. 김태용. *인공지능을 활용한 사회과학 연구방법*. 서울: 교보문고, 2023. https://product.kyobobook.co.kr/detail/S000213184420.

Kirschenbaum, Matthew. "Prepare for the Textpocalypse." *The Atlantic*, March 8, 2023. https://www.theatlantic.com/technology/archive/2023/03/ai-chatgpt-writing-language-models/673318/.

Kirschenbaum, Matthew, and Rita Raley. "AI and the University as a Service." *PMLA/Publications of the Modern Language Association of America* 139, no. 3 (2024): 504–15. https://doi.org/10.1632/S003081292400052X.

Knapp, Steven, and Walter Benn Michaels. "Against Theory." *Critical Inquiry* 8, no. 4 (1982): 723–742. https://doi.org/10.1086/448178.

Eisenstein, Jacob. *Introduction to Natural Language Processing*. Cambridge, MA: MIT Press, 2019. https://lccn.loc.gov/2018059552.

Eun-Young Kwon. 2024. "Analysis of trends in domestic mobile-assisted language learning in English education: Using text mining." Korea Journal of English Language and Linguistics 24: 325-347. 10.15738/kjell.24..202404.325. 권은영.

2024. "국내 모바일 기반 영어교육 연구 동향분석: 텍스트마이닝을 활용하여." 영어학 24: 325-347. 10.15738/kjell.24..202404.325

Ishiguro, Kazuo. Klara and the Sun: A Novel. New York: Alfred A. Knopf, 2021. https://lccn.loc.gov/2020015327.

Laquintano, Timothy, and Annette Vee. "AI and the Everyday Writer." *PMLA/Publications of the Modern Language Association of America* 139, no. 3 (2024): 527–32. https://doi.org/10.1632/S0030812924000609.

Lee, Min-Chul. "Kiwi: Developing a Korean Morphological Analyzer Based on Statistical Language Models and Skip-Bigram." *Korean Journal of Digital Humanities* 1, no. 1 (2023): 109–135. https://accesson.kr/kjdh/v.1/1/109/43508. 이민철, "Kiwi: 통계적 언어 모델과 Skip-Bigram 을 이용한 한국어 형태소 분석기 구현." *디지털인문학* 1, no. 1 (May. 2024): 109-136, https://doi.org/10.23287/KJDH.2024.1.1.6.

Li, Z., and Volker Hegelheimer. "Mobile-Assisted Grammar Exercises: Effects on Self-Editing in L2 Writing." *Language Learning & Technology* 17, no. 3 (2013): 135–156.

Maeil Business Newspaper, "Announcement of the Technical Report emphasis on multilingual skills "Prove global competitiveness" *Maeil Business Newspaper*, April 4, 2024, https://www.mk.co.kr/en/it/10982237

Martin, Meredith. "Command Lines for the Humanities." *PMLA/Publications of the Modern Language Association of America* 139, no. 3 (2024): 541–47. https://doi.org/10.1632/S0030812924000555.

Muñoz-Ortiz, Alberto, Carlos Gómez-Rodríguez, and David Vilares. "Contrasting Linguistic Patterns in Human and LLM-Generated News Text." *Artificial Intelligence Review* 57 (2024): 265–292. https://doi.org/10.1007/s10462-024-10903-2.

NAVER Cloud, HyperCLOVA X Team, "HyperCLOVA X Technical Report," *arXiv*, April 2024, https://arxiv.org/abs/2404.01954.

Perlow, Seth. "Generative Theories, Pretrained Responses: Large AI Models and the Humanities." *PMLA/Publications of the Modern Language Association of America* 139, no. 3 (2024): 548–52. https://doi.org/10.1632/S0030812924000518.

PMLA, "Special Topic: Varieties of Digital Humanities," vol. 135, no. 1 (January 2020). https://www.cambridge.org/core/journals/pmla/issue/847B9F9004B8462A86F0323153F760B1

Sajid, Haziqa. "The State of Multilingual LLMs: Moving Beyond English." *Unite.AI*, February 10, 2024. https://www.unite.ai/the-state-of-multilingual-llms-moving-beyond-english/.

Shoemaker, Tyler. "Machines, Reading," *Critical Inquiry* (blog), June 28, 2023, https://critinq.wordpress.com/2023/06/28/machines-reading/.

Slater, Avery. "Hermenautics: Toward a Disinformation Theory." *New Literary History* 54, no. 2 (2023): 1257–1261. https://doi.org/10.1353/nlh.2023.a907171.

Taecharungroj, Viriya. "What Can ChatGPT Do? Analyzing Early Reactions to the Innovative AI Chatbot on Twitter." *Big Data and Cognitive Computing* 7, no. 1 (2023): 35–45.

Vadde, Aarthi. "Inside and Outside the Language Machines." *PMLA/Publications of the Modern Language Association of America* 139, no. 3 (2024): 553–58. https://doi.org/10.1632/S0030812924000579.

Yoo, Wonjun and Sang, Jun, *Introduction to Natural Language Processing Using Deep Learning*, 2024, accessed December 13, 2024, https://wikidocs.net/book/2155.

Zadorozhnyy, Artem, and Wan Yee Winsy Lai. 2024. "ChatGPT and L2 Written Communication: A Game-Changer or Just Another Tool?" *Languages* 9, no. 1: 5. https://doi.org/10.3390/languages9010005