

# No Longer Only Human: Language in the Age of AI

Tom GALLY<sup>1</sup>

Language educators teach based on theories of language. Whether choosing materials, preparing lesson plans, conducting classes, or planning curriculums, each teacher decides what to do and how to teach based in part on an abstract framework of concepts, relationships, hypotheses, and claims about the subject they teach—in other words, a theory of language. That theory of language might be formal—explicated in scholarly books and research papers, taught to novice teachers—or it might be informal, a collection of ideas and beliefs that the teacher has acquired through experience.

Different theories, whether formal or informal, approach language in different ways. Some theories, and consequently some teachers, regard language as primarily a cognitive phenomenon, as a system of sounds, words, grammar, and meaning processed in the language user's brain and mind. Others focus more on language as a medium of one-to-one communication, of conversation and dialog. And still others are concerned most with the role of language in society: its dialects and variations; its associations with identity, class, gender, ideology, and power; its role as a transmitter and shaper of culture. But regardless of their formality and focus, all theories of natural language until now have, explicitly or implicitly, been grounded on one assumption: that language is something only human beings possess and that we use language only among ourselves.

---

<sup>1</sup> Emeritus Professor, The University of Tokyo. Formerly affiliated with the Department of English Language and Department of Language and Information Sciences. Currently Project Professor, Center for Global Education.

That assumption is challenged by recent advances in artificial intelligence. Now we can use natural language to interact not only with each other but with computers as well. Using the same words and sentences we would use with another person, we can ask computers questions, engage them in conversation, tell them to draw pictures and write computer programs—even have them teach us other languages. These advances in AI compel us to reconsider our traditional views about human-only conceptions of language and about how natural languages can be taught and learned. In this new age of AI, language educators need new theories of language.

### A Turning Point

On November 30, 2022, the American company OpenAI released to the public an interactive chatbot called ChatGPT. As I wrote a few weeks later, in an essay published in this journal in March 2023, ChatGPT amazed me with its language ability. It could respond appropriately to written prompts, correctly identify the meanings of words in context, compose well-organized essays, and, especially, converse with me in a back-and-forth dialogue almost as if it were human. Unlike programming languages, it was forgiving of sloppiness: it responded appropriately to prompts that contained typos, grammar mistakes, and nonstandard expressions, and it could guess correctly the intended meaning of ambiguous statements. Asking it questions, telling it things, and reading its responses felt like interacting with another person.

Interest in ChatGPT and other forms of generative AI spread quickly beyond early adopters like me. Within a few months of its release, schools and universities became aware that their students could use it to do homework, to write papers, and to cheat, and they began issuing guidelines for students' use of AI. Mainstream media ran many reports and commentary on this new development, and social media filled with examples of ChatGPT's output and discussions of whether it was really intelligent or not. Warned of looming dangers from superhuman AI, governments began to discuss the regulation of AI technology.

Technological progress also sped up. In March 2023, OpenAI released a more powerful and reliable model called GPT-4. Later

in the year, that model would be upgraded again, acquiring multimodal capabilities, including a spoken interface and image recognition and creation functions. Competing models from Google, Anthropic, Meta, Amazon, Baidu, and other commercial and open-source developers were also announced. Auxiliary technology to support language use with computers improved as well: speech recognition became more accurate, and synthetic voices went from sounding robotic to being indistinguishable from humans'. While some people predicted that AI progress was about to plateau, many observers—including me—felt that a turning point had been reached, and that the coming months and years would see even faster improvements in artificial intelligence of all types.

In my previous essay, I predicted that a new era was arriving for education, and my prediction seems to have been largely correct. Hundreds of books, videos, and online guides about how to use ChatGPT for learning languages and other subjects have appeared, and smartphone app stores have filled up with AI-powered conversation tools. School and university educators have been more cautious about the technology, but by late 2023 reports of teachers incorporating generative AI in their lessons were becoming common. Even educational administrators and government policymakers—whom I had predicted would be slow to grasp and respond to this challenge—recognized the potential and dangers of AI, and initiatives to integrate AI with school and university education were under way in many countries.

## Challenges

While AI's impact began to be felt in all areas of teaching and learning, it was perhaps in language education where the possibilities and threats were strongest. After all, the breakthrough systems were large *language* models, and their greatest strength, especially at first, was interacting with people using natural language. And the biggest weakness of the first LLMs—their tendency toward hallucination or confabulation, that is, the blithe production of falsehoods—was not so serious for language education. While people trying to acquire explicit knowledge about a subject need to be cautious about what LLMs tell them, the

goal of language learning is often more to acquire familiarity and fluency with linguistic forms. In such cases, occasional falsehoods may not be a serious problem. After all, when practicing using a second language with a human partner, one might be annoyed if the other person makes up facts occasionally, but those confabulations may not interfere much with the language practice itself.

I should note that the first publicly available LLMs did exhibit a quirk that was worrisome for language learning: a weakness at describing language itself. The models were trained on texts that had been converted into units called tokens, which did not always correspond one-on-one to letters, characters, morphemes, words, or other units that humans normally use. While the first release of ChatGPT was very good at explaining the meanings of texts, because it couldn't "see" the individual components of sentences it sometimes gave false explanations of grammar and other linguistic structures. GPT-4 was more accurate though still not perfect, while other LLMs could not even count reliably the number of words in a sentence. LLMs exhibited similar weaknesses with logic, mathematics, and everyday common sense.

In that first year, there were other barriers to the immediate adoption of generative AI for language learning. One was the cost. While a free version of ChatGPT was available, access to OpenAI's best model required a subscription of 20 U.S. dollars a month—more than many schools and universities could ask their students to pay. Another was that the techniques for using the models—sometimes called "prompt engineering"—were still in their infancy, and new users were often unable to use LLMs effectively. There were also concerns about privacy. Generative AI models that could be securely hosted on a personal computer were significantly weaker than ChatGPT; the only way to interact with the strongest models was to use cloud-based services. And many teachers, and some students, were uncomfortable about incorporating AI into language learning. They felt that language education was fundamentally a human endeavor and that allowing AI to play a central role in learning might harm students' language acquisition and motivation to continue.

Despite the drawbacks and resistance and despite uncertainty about AI's future, the performance, flexibility, and ease of use of generative AI seem to ensure that it will play an increasingly

important role in language education in the years to come. But this prospect raises the profound issue mentioned earlier. Never before in human history has it been possible to communicate using natural language with anything other than humans. What does it mean for our conceptions of language—our linguistic theories—and for our ways of teaching and learning languages that humans now also use our languages to interact with computers?

## Sounds

For some aspects of language, the impact of AI may be slight. One example is articulatory phonetics, that is, the study of how people make speech sounds by passing air from their lungs through their throats, vocal cords, oral and nasal cavities, and lips. Students learning how to produce the sounds of another language will still need to imitate the physical movements of human speakers of that language. Because computers make sounds by transmitting electrical currents through speakers, there would be no point in trying to imitate their speech production process.

Another area not yet challenged by AI is gestures—hand motions, facial expressions, and head shaking, which can be different for different languages. Current AI models recognize only what is typed or spoken by humans and cannot perceive or display gestures. Until onscreen characters or humanoid robots that recognize and produce gestures when interacting with humans become available, we may not need to readjust how we think about gestures as a part of language.

The theoretical implications are a bit more significant for acoustic phonetics, that is, the study of the physical properties of speech sounds transmitted through the air. When two humans have a spoken conversation, the sounds of one person's speech reach the other person's ears, where they are processed by the listener's auditory nerves and brain. In contrast, when a language-capable computer receives speech sounds, a microphone converts the sounds into analog signals, which are then converted to digital form for processing—a method very different from the human one. Previously, language students trying to speak a second language had to think only about the impression

their speech sounds made on other people and whether their listeners understood them. Now, however, a student practicing pronunciation with a computer will get feedback based on how the AI-equipped computer receives and processes the sounds. Approaches to teaching and learning pronunciation will need to be rethought.

According to modern theories of language, humans perceive the sounds of speech not in terms of the air vibrations that reach our ears. Rather, we interpret those sounds as higher-level linguistic units called phonemes, which correspond only partially to actual sonic waves. The teaching of second-language pronunciation and listening comprehension thus focuses on phonemes. Unlike the physical sounds of speech, which vary greatly with each utterance, phonemes are largely constant from utterance to utterance and from speaker to speaker. While phoneme-based speech recognition systems have been developed, today's more successful systems use deep learning and neural networks to convert audio waveforms into a numerical transcription that does not have an explicit phonemic representation. A theory of phonology that includes interaction between humans and AI must therefore consider the very different ways humans and computers process the sounds of language. Both humans' recognition of AI-produced speech and the AI's recognition of human speech are likely to be significantly different from what happens in human-to-human conversation.

## Grammar

When we move from the sounds of language to the more abstract level of grammar, one might think that LLMs process language similarly to human beings. After all, theories of grammar have often been based on fairly rigid concepts—parts of speech, inflectional categories, syntactic functions, etc.—and on deterministic, almost mathematical rules. Some theories of generative grammar have even resembled computer programs.

But LLMs acquire and process the grammar of natural languages in a way fundamentally different from human beings. According to the theories of Noam Chomsky and others, humans are thought to have an innate general language function that becomes specialized on a specific language through limited

exposure to and use of that language during early childhood. In contrast, LLMs begin their training without any preconceptions about linguistic structure, and they require vastly more linguistic input than any child receives in order to become able to interact in a language convincingly. Through a method known as unsupervised learning, LLMs are fed large corpora of text, from which they derive numerical representations that can, after much processing, produce linguistic patterns and structures similar to humans'. An LLM's grammatical "understanding" is thus not rule-based but statistical and probabilistic. While a student of a language might be taught that a past-tense verb must agree with its subject in number, person, and gender and while a child acquiring that language can apply that rule after exposure to perhaps a few dozen examples, an LLM becomes able to conjugate the verb correctly only after processing thousands of examples of such agreement. How these wide differences in methods of acquiring and processing grammar might affect human-AI interaction is not yet clear. But the ability of LLMs to produce grammatically correct sentences without anything that corresponds to the human predisposition to language acquisition suggests that some current theories of grammar need to be rethought.

## Nativeness

One area where the different approaches to grammar by humans and AI might be especially important is native language competence. While some aspects of native language acquisition and of social attitudes about native speakers are controversial, it is largely agreed that, in general, people who acquire a language as children have a qualitatively different ability from those who learn the language as adults. People who grow up speaking a language can process certain aspects of grammar—tense and aspect, subordination, pronoun reference, definiteness and indefiniteness—correctly, fluently, and unconsciously, while those who start later usually never acquire full competence. Whatever mechanism in the human brain might cause this change during the transition from childhood to adulthood, it seems to have no counterpart in LLMs. Statistical processing of text data does not change as the computer's components age,

and all that the models need to imitate native grammatical competence convincingly is a sufficiently large corpus of text that reflects the performance of native speakers and sufficient processing capacity. To the extent that theories of language and language education have incorporated the concept of nativeness, they will now need to reflect the fact that nativeness does not apply to the LLMs that people are increasingly interacting with.

## Meaning

While theories of the sounds and grammar of language will need some revision for the age of AI, perhaps the two aspects of language that will require the most rethinking are semantics and pragmatics, that is, the representation and conveyance of meaning by language and the use of language for practical purposes.

Linguistic meaning has been theorized in various ways. Informally, the meanings of words or phrases are often seen as corresponding directly to concrete or abstract entities. This denotative approach is common in second-language education, in which words may be taught and learned as though they had simple, fixed meanings, like synonyms in a glossary. For beginning learners, that simplistic approach to lexical meaning might be unavoidable, but it soon becomes untenable when learners try to grasp the meanings of words in actual use. In real-life situations, the associative and emotional connotations of words and the subtle, context-dependent changes in meaning make a purely denotative approach to meaning inadequate.

Perhaps the most important—and also most complex—frameworks for thinking about meaning focus on how language is used in interpersonal interactions and other social contexts. Philosophers of language with a social bent have theorized meaning in terms of what people do using language, how people co-create meaning when talking with each other, and how social status, gender, ideologies, and power relations affect how meaning forms and changes. In such approaches, language is not a symbolic signaling system comparable to a semaphore, telegraph, or HTTP connection. Rather, it is a dynamic, fluid, multifaceted medium that is created by and dependent on the people who speak it and the situations in which it is used. And, once



again, a fundamental assumption of all such approaches is that the users of language are people and only people.

In social theories of language, each language user has various statuses and roles in society: as a person with a name, age, gender identity, nationality, family, cultural background, and occupation, as someone with likes and dislikes and desires and goals. All of those factors affect how people use language to interact with each other. And what we say is influenced not only by our own personal characteristics; we also change how we talk based on what we believe about the people we are talking with. One of the most challenging parts of becoming proficient in a second language is learning what linguistic forms and conversational strategies to use when addressing others, when asking questions and making requests, when talking about oneself and others. One must not only be able to choose the correct words and sentence patterns; one must also infer what kind of person the listener is and how they might react. We can ignore such factors when interacting with AI, for it has no social status, no personal identity, no individual will, and it will not get annoyed or angered by a learner's ignorance or rudeness. Our conceptions of language as a social instrument must therefore change now that these new participants in our linguistic interactions are not members of our human society.

A related issue is psychological: When people speak with each other, they try to infer, consciously or unconsciously, the mental and emotional state of their interlocutor. Does the other person know what I am referring to? Is the other person interested in what I am saying? What does the other person want? Is the other person pleased, angry, or bored? Our inferences affect how we use language in very concrete ways. In English, for example, we use pronouns, demonstratives, and the definite article when we assume that the other person knows what we are referring to, and we vary the forms of requests based on how we think the listener will react. How must the theorizing and teaching of such linguistic forms change when people are increasingly using them to interact with entities that have no consciousness, emotions, or motivations?

## Culture

In recent decades, the teaching of English as a foreign language in Japan has gradually de-emphasized the role of culture in language. Awareness of English's growing importance as a widely spoken second language and international lingua franca has made English education's traditional focus on learning about British and American culture seem misguided. However, for other languages—Korean, German, Mandarin Chinese—the association between learning the language and learning about the culture of the people who speak it is still strong. And while culture plays less of a role in school and university English education, individual learners still find value in studying the language through the cultures of English-speaking countries, especially if they plan to visit or live in those countries or they are interested in some aspect of their particular cultures.

The language models now available are largely acultural. Although they have been developed by U.S. companies, have ethical constraints imposed on them that reflect a certain subset of American values, and use American English as their default output, the bots do not really embody American or any other cultural background. Much discussion has focused on how biases in their training data can lead to biases in their output, but an LLM's biases are very different from any individual human's. An important part of learning to interact with people in a foreign language is becoming aware of how their knowledge, assumptions, and predilections are different from one's own. The knowledge, assumptions, and predilections exhibited by LLMs are the result of the massive data they were trained on, not of having been raised and educated as a human member of a particular human culture.

## New Theories

A theory does not need to be a rigorous academic framework with explicit definitions, claims, and methodologies. Informal theories, personal systems of understanding, are also useful—indeed, they are unavoidable. Given the complexities of the world in which we live, we have no choice but to use abstract concepts and principles to make sense of what we experience.

Language in particular would be difficult to think about, and probably impossible to teach, without at least an informal theory. Even a language teacher with no academic training thinks in terms of abstract linguistic concepts: word, sentence, pronunciation, accent, politeness, clarity, correctness. Though perhaps not fully rigorous or internally consistent, those concepts and the teacher's understanding of them constitute that teacher's theory of language. Without such a theory, the teacher would be unable to teach.

At some point, human interaction with AI will have to be incorporated into formal theories of human language, theories that could form the frameworks for rigorous academic research. But technology is evolving too rapidly to make that feasible now. Any full-fledged theory proposed today could easily go out of date as the AI becomes more capable and multimodal and as it acquires more human-like characteristics, such as consciousness, embodiment, personality, and active roles in society.

But language educators cannot wait for the technology to mature. Our students are already using AI to learn and practice language, and teachers need to figure out how best to guide them. We need new theories of language now, new ways of understanding how people use human language while interacting with today's intelligent-seeming, linguistically fluent, but nevertheless unhuman bots. Because language teachers usually draw on their own experience learning languages, they need to interact with the latest bots themselves to gain a sense of what it is like to learn a language from them. They should also observe their students using the AI and ask them for their impressions. They should try to figure out how interacting with AI will shape their students' learning and language acquisition and how it might motivate or demotivate them. Most of all, they should be willing to modify their long-held beliefs about what language is and how it is learned—that is, their theories of language. Since the arrival of these amazing new language models, these fluent, engaging, and sometimes frightening nonhuman conversation partners, our old ways of thinking about language need to change.