

GEOFF VAUGHAN – RITA SZÚTS-NOVÁK

WHAT WE TALK ABOUT WHEN WE TALK ABOUT AI. ARTIFICIAL
INTELLIGENCE AND LANGUAGE

Abstract

With the introduction of AI tools and software into an increasingly large number of applications that are used across society, including education, it is of significant importance to consider the nature, role and use of language, both in terms of the language at the heart of large language models and the language at the heart of humans. Language has been the evolutionary miracle that has set humans apart from the rest of the animal kingdom. It has led to the production of language in the form of written records and spoken output, which has shaped our cultural and social life. The use of AI large language models and tools has now given us the potential to regenerate a human interest in what Merlin Donald referred to as “our collective memory palace”. It may bring us closer to a deeper, linguistic appreciation of the riches of humanity’s creative and intellectual treasures. This article takes a brief look at some critical areas of language history and theory. It identifies the importance of the library as a crucial public institution that can help shape the relationship between AI and humans.

Keywords: artificial intelligence, large language models, language, humanity, libraries, language theory

Background

The initial, global buzz around generative AI tools that were released in November 2022 was the excitement of having instant linguistic interactions with a screen. Here was a brave new world where questions were typed into a simple display panel, and an impressive array of text instantly flowed back on the screen. To many, the experience was one of amazement often followed by a desire to immediately discover the extent that interactions could go to. Here we were, interacting with a machine that talked, or messaged, back to us and communicated in such a way that we had never experienced before. A human-to-computer language interface. It was not long before many discovered the microphone icon embedded on the interface page, and the interaction with the AI program became a two-way voice dialogue. Not only could the AI voice give impressive responses, but it could easily be asked to reframe its responsive style and even the language used. Ask a question in English, receive an answer, even in another language, and adapt the style

and voice register at will. Here we were in a new linguistic paradigm with technology. At the very heart of the new generative AI tools, or LLMs, is the word language. The importance of language and the importance of our understanding of these new forms of technology with regards to language are extremely significant as these tools become more present in our everyday lives. In this article, we look at some important issues of language and make an individual case for seeing how our libraries and their potential use of AI as an area of significant importance in this relationship may develop (Kókuti et al. 2023; Rajcsányi-Molnár et al. 2024a, 2024b).

Linguistic knowledge

What exactly can we do as we are faced with the inextricable rise of AI and the inevitable challenges to everything that we have previously considered at least academically stable with regards to language. One crucial area of consideration is language use. In linguistics, the field of pragmatism is concerned with the 'use' of language; what Saussure termed 'parole'. This external use of language, or surface structure, in linguistics has a far more settled contemporary understanding than the more complex and challenging linguistic study of deep structures of language and is neatly summarised by Berwick and Chomsky thus:

“[...]certain key biological components associated with language, in particular the 'input-output' system of vocal learning and production that constitutes part of the system we will call 'externalizaion', have been clarified biologically and genetically, so much so that we can effectively use a 'divide-and-conquer' strategy and place this sensorimotor aspect of externalization aside while we focus on language's more central properties” (Berwick and Chomsky 2011: 2).

Although of crucial relevance to the understanding of how language actually is formed in the human mind, issues of acquisition and deep structure can generally be put to one side when looking at the pragmatics of language; that is, our everyday use of language and what that may have to say about our relationship with generative technology and language.

It is the overwhelming and fundamental success of homo sapiens that we have developed a language ability that has allowed us to convey a thought or idea from one person to another or many. The evolution of human language shows that this capacity for using language to convey a thought or an idea has been a long and hard-won process. Estimates for the human ability to

speak range between 100,000 and 50,000 years ago. It is the unique human capacity, and it has been a process of change and refinement ever since. It is also a uniquely social process. As Tomasello (2011: 343) explains:

“Language, or better linguistic communication, is (thus) not any kind of object, formal or otherwise; rather it is a form of social action constituted by social conventions for achieving social ends, premised on at least some shared understanding and shared purposes among users.”

Humans are bombarded with language from the minute they are born. A process of symbolic interpretation of spoken language starts in the early months of a new-born child’s life, and the growth in linguistic comprehension and production is one of the most incredible human cognitive achievements that truly separates us from the rest of the animal kingdom. Utterances begin to occur as early as eight months, and a fully developed spoken syntax generally forms in a child by the age of three. Young humans are truly remarkable language creators. Dehaene (2019: 67) explains this incredible human linguistic feat as:

“In a few months, children quickly manage to surpass any existing artificial intelligence algorithm. By the time they blow out the candle, they have already laid down the foundation for the main rules of their native language at several levels, from elementary sounds (phonemes) to melody (prosody), vocabulary (lexicon), and grammar rules (syntax).”

Becoming literate

Our ability to represent speech with visual or written symbols is even more recent in terms of evolutionary time; our first written symbols appeared approximately 5000 years ago. Again, the development and refinement of the written word has been a long, hard-won process. Our biologically primary spoken capacity for language and our non-biological secondary ability to use a written symbolic code for writing are uniquely human attributes. No other animal can do anything remotely what we do in terms of speech and symbolic representation.

The technological invention of the printing press in the 15th century dramatically transformed our human relationship with text. From that point onward, written language, which had been the preserve of the elite, often religious, scribes, gradually allowed the general population to become literate. Again, this has been a hard-won transformation that is still being

played out in education systems across the globe, as can be seen in the huge efforts that are poured into elementary education in order to achieve a level of functional literacy.

The education system is infused with language both from without and within. We are inextricably tied to language, as is evidenced by the insistence on learning the foundations of written language and literacy in most elementary stages of education globally. Donald (2001a: 232) explains this as a process which eventually becomes automatic, and

“once they are learned, the demons will go downstairs, the ideograms will become invisible, processed down below, in the automatized basement of the mind, and your attention will focus only on the meaning of the text”.

It is this immense cognitive battle that has been at the heart of education, arguably since classical times for some and certainly since the inception of mass education in the nineteenth century. Donald makes the important point that once basic decoding of an alphabet has been mastered the attention turns to meaning. Quine (1968: 143) explains this crucial point thus:

“There are two parts to knowing a word. One part is being familiar with the sound of it and being able to reproduce it... The other part, the semantic part, is knowing how to use the word.”

Developing and mastering language use is at the core and heart of schooling. From the initial teaching of the ability to decode the alphabet through to a functioning literacy by the end of the elementary schooling period students are subjected to a relentless focus on language. The relentless addiction to written signs and symbols continues to be the major element of education in the upper years. Students are asked to read increasingly complex text, asked to answer questions that assess their comprehension and appreciation of those increasingly complex texts and asked to become adept at producing a wider range of written texts in the form of essays, arguments, persuasive texts, among a few of the tasks that students are faced with as they progress through the higher stages of the formal schooling stage. What is being increasingly asked of students is that they attend to a precision of language. As Austin (1976: 73) says:

“Language as such and in its primitive stage is not precise, and it is also not, in our sense, explicit: precision in language makes it clearer what is being said – its meaning: explicitness, in our sense, makes clearer the force of the utterances, or 'how it is to be taken'.”

By the time students graduate from secondary education, many are regarded as having mastered a level of linguistic ability that, when viewed against the timeline of human cognitive evolution, is truly staggering. True, some students have not been able to reach that level of mastery and the failures of mature education systems with regards to this are constantly being reviewed by many societal organisations, from education schools, governments and global agencies. But let us take a positive, holistic view of the literacy ability of global young learners today. By any measure they are generally extremely literate having navigated the journey of language over their first eighteen years. Despite regular dire warnings as to the current state of literacy, from a wider perspective, the transformation of literacy for the majority of young people worldwide has been one of the most significant social achievements in history.

The whole biosphere of further academic development has traditionally relied on the key role of language, and the development of the ability for learners to improve continually has been a constant in the fields of psychology, learning science and most other academic disciplines. Admittedly, our "collective memory banks" (Donald 2001b: 569), the traditional storehouses of knowledge such as books, scientific papers, articles and journals, have multiplied to an almost infinite size in the digital age and have left us at the stage where the most pressing problem has become that of "how can we educate people to handle the sheer immensity of our global memory palace?"

The introduction of artificial intelligence

The introduction of AI large language models has further challenged the nature of our relationship between our human ability to navigate a collective language space to an almost impossible level. Today, this whole ecosystem has come face to face with an existential crisis of meaning. The control of language has moved from institution to device. The role of the expert has moved from the classroom to the personal screen and headphones. The gatekeepers of knowledge, with language at the heart of the process, have shifted from the expert to the machine. And this has profound consequences for all. Bruner (1996: 7) saw this dichotomy between a cultural and computational view of meaning-making: There is no decision procedure known that could resolve the question of whether the incommensurability between culturalism's meaning-making and computational's information processing could ever be overcome. Yet, for all that, the two have a kinship that is difficult to ignore.

Once meanings are established, it is their formalizations into a well-formed category system that computational rules can manage.

It is of interest that Bruner chose the term 'kinship' to describe the possibility of combining our cultural and computational meaning-making systems or what Donald (1991: 358) has described as our "hybrid minds".

Our new, computational-based, artificial language models fundamentally shift our relationship with language. These models force us to confront the relationship between language and meaning. As Clark (2024) in his concluding paragraph states: It turns out that this technology is not about technology-in-itself, something out there to be tamed. It is about US! Our collective cultural legacy has been used to train these amazing models. It is we speaking to ourselves, asking what we should do next. The Copernican revolution is not something out there but within ourselves. We are back at the centre of our relationship with knowledge and our future.

Importantly, Clark highlights these large language models' transactional, dialogic nature. The interactivity we can now enjoy with being able to interrogate and receive instant feedback with these models is bringing us much closer to what has been our experience of what has been the social development of language that we have experienced as humans throughout our history. As Tomasello stressed, we learn through linguistic interaction. However, we do need to remain acutely aware of our human need for quality control and our capacity for judgement of language use. For it was Montaigne, writing in the second half of the sixteenth century, who warned us that:

"Just as our mind is strengthened by communication with vigorous and orderly intellects, it is impossible to say how much it loses and is debased by our constant intercourse and association with mean and feeble intellects" (Clark 2024: 287).

As we begin to interact and 'talk' with AI in a more natural linguistic communicative way we have to recognise that our interactions are with a computational model that, according to Donald (1991: 24) "knows nothing of the world to which it refers".

A fundamental shift and the future role of libraries

Today we are witnessing a fundamental shift in our relationship between human language and our digital technologies. We are witnessing the power of both AI and the power of language and it is increasingly returning to the original functions of language – of speech and thought. The challenge that

we face today is how to find some practical, yet innovative, ways to connect our age-old human quest for knowledge and understanding through language with these new language-based machines. These are not easy challenges and, as we progress with understanding our relationship with artificial intelligence, will require expertise in what Searle (2008: 10) referred to as the “future of philosophy” in such areas as the philosophy of language and of society (Balázs 2023a, b).

It is our contention that we do indeed have a vital and long-standing, structural and intellectual model in our libraries that can help us navigate this changing world. In Donald’s (2001a: 572) evaluation of our relationship with our collected and stored knowledge, he states the fundamental question that we are facing:

“A growing mind still has the same basic need for structure as it ever had. Cultural astronauts need a home planet before venturing into representational space. They need a map and a guide, at the very least, and some rules.”

The challenges of the information age and technological developments, including the advances in AI technology, have created a new perspective for public collections in content collection, storage navigation and interactivity (Arató–Balázs 2023: 756). The traditional museum and library practice approach is gradually changing in line with the expectations of the new digital age. Changing attitudes is of paramount importance in involving users. Public collections have traditionally been institutions dedicated to the preservation and transfer of information. However, the rise of digital technologies and the spread of the internet have opened the door to interactivity and community collaboration (Balázs 2023b). In this new approach, public collections are no longer only places of storage and research, but also active participants in knowledge sharing and communal, linguistic creativity. With the help of large language models, digital tools, visitors, researchers and interested parties will be able to more actively participate in connecting more closely with the content of library connections. The memory palaces will hopefully become palaces of active, fruitful and important discussion.

This change of attitude will not only increase the diversity of the content of public collections, but will also enhance the experience of involving users. Interactive public collections not only offer static information, but also provide opportunities for active dialogue and cooperation between audiences and institutions. The resulting public collection content will become much more relevant and authentic for users, as community members become part

of preserving knowledge and culture. In addition, public collections built in this way are fit for the modern digital age and offer new opportunities for education, research and creativity. Researchers can collaborate on their discoveries through virtual networks. Virtual data and knowledge networks also reveal previously unexplored connections, while ever-expanding connections provide new resources for research. The growing wealth of data and information provides new inputs for AI research and development and creates synergies for growth that can be exploited in both processes. Ultimately, by using large language models in connecting with the content of the treasures from our collective, stored records in our libraries we would hope for a regeneration of interest in the linguistic riches that have always been there waiting to be found by the reading public (Balázs 2021: 56)

Conclusion

There is a need for educators, and all those associated with education in wider society, to begin to focus on the role and importance of human language when dealing with this new world of AI. For many in our new age of computation, the actual human processes of language acquisition and the qualities of language that make us the truly unique species have often been assumed. In many ways, it has been taken for granted that a biologically primary function (that of human speech) has been enough. It is by no means enough. The daunting pedagogical task of achieving symbolic literacy that every one of us must go through and that is the beginning of an incredibly complex marriage of both cognition and culture in our socio-cultural worlds is one that we should begin to appreciate in this new age, ultimately because it is of such vital importance to our understanding of ourselves as cognitive beings as Sperber and Hirschfeld (1999: 127) express: “Ordinary cognitive activity does not take place in a fixed experimental setting where the information is available is strictly limited and controlled, but in a complex, information-rich, ever-changing environment... In the human case, moreover, the environment is densely furnished with cultural objects and events most of which have, at least in part, the function of producing cognitive effects.”

It is our hope that, as we navigate the future of artificial intelligence use, we may reach a reconnection with the wonders and mysteries of human language which have been the subject of an incredible process of cultural development over thousands of years, where we often find ourselves stunned by the immense power of words to express what sometimes can not even be expressed - the innermost workings of our human minds. What lies hidden beneath the surface of language, and AI, may be exactly where the truth may

be hiding. As E.M. Forster (1926: 93) in a stunning moment of insight in *A Passage to India* put so very well: “Mrs Moore shivered, ‘A ghost!’ But the idea of a ghost scarcely passed her lips. The young people did not take it up, being occupied with their own outlooks, and deprived of support it perished, or was reabsorbed into the part of the mind that seldom speaks.”

References

- Arató, Balázs – Balázs, Géza 2023. Hol a tudás, amit elvesztettünk az információban? A digitális pedagógia diszkurzív-intuitív megközelítése antropológiai és jogi nézőpontból. *Magyar Nyelvőr*, 2023: 755–76.
<https://doi.org/10.38143/Nyr.2023.5.755>
- Austin, John Langshaw 1976. *How To Do Things With Words*, second edition, Oxford University Press, London.
<https://doi.org/10.1093/acprof:oso/9780198245537.0010001>
- Balázs, Géza 2023a. Tamás Ildikó: „Adj netet!” *Magyar Nyelvőr*, 147. 2023: 120–7.
<https://doi.org/10.38143/Nyr.2023.1.120>
- Balázs, Géza 2023b. *Az internet népe. Internet – társadalom – kultúra – nyelv. A kulturális és a tervezett evolúció határán.* Ludovika Egyetemi Kiadó, Budapest.
- Balázs, László – Rajcsányi-Molnár, Mónika – András, István – Sitku, Krisztina 2021. Social Responsibility and Community Engagement at a Hungarian Regional University. *Journal of Higher Education Theory and Practice*, 21. 54–63.
- Berwick, Robert C. – Chomsky, Noam 2016. *Why Only Us: Language and Evolution.* The MIT Press, Cambridge.
<https://doi.org/10.7551/mitpress/9780262034241.0010001>
- Bruner, Jerome 1996. *The Culture of Education.* Harvard University Press. Cambridge. <https://doi.org/10.4159/9780674251083>
- Clark, Donald 2024. *AI is the Copernican collective, cultural mind - it is US!* <https://donaldclarkplanb.blogspot.com/2024/02/ai-is-copernican-collective-cultural.html>
- Dehaene, Stanislas 2021. *How We Learn: The New Science of Education and the Brain.* Penguin Books, New York.
- Donald, Merlin 1991. *Origins of the Modern Mind.* Harvard University Press. Cambridge.
- Donald, Merlin 2001a. *A Mind So Rare.* Norton. New York.
- Donald, Merlin 2001b. Memory Palaces: The Revolutionary Function of Libraries. *Queen's Quarterly*, 108(4) 559–72.
- Forster, Edward Morgan 1965. *A Passage to India.* Progress Publishers. Moscow.
- Kókuti, Tamás – Balázs, László – András, István – Rajcsányi-Molnár, Mónika 2023. Collaborating with Artificial Intelligence – AI in Business Communication Education. In: *IEEE - Óbudai Egyetem; IEEE (szerk.) 2023 IEEE 6th International Conference and Workshop Óbuda on Electrical and Power Engineering (CANDO-EPE): Proceedings Danvers (MA), IEEE.* 287–94.

- Montaigne 1958. *Essays*. Penguin Books. New York.
- Rajcsányi-Molnár, Mónika – Balázs, László – András, István – Czifra, Sándor 2024a. Competition as an Effective Motivational Tool in Online Education. In: Anikó, Szakál (ed.): *IEEE 11th International Conference on Computational Cybernetics and Cyber-Medical Systems: ICC 2024: Proceedings*. IEEE Hungary Section Budapest. 83–8.
- Rajcsányi-Molnár, Mónika – Balázs, László – András, István 2024b. Online Leadership Training in Higher Education Environment. *ACTA POLYTECHNICA HUNGARICA* 21, 39–52.
- Quine, Willard van Orman 1968. Ontological Reality. *Journal of Philosophy*, 65(7) 185–206.
- Searle, J. 2008. *Philosophy in a New Century: Selected Essays*. Cambridge University Press, UK. <https://doi.org/10.1017/CBO9780511812859>
- Sperber, Dan and Hirschfield, Lawrence 1999. Culture, Cognition and Evolution. In: Robert A. Wilson and Frank C. Keil (eds) *MIT Encyclopedia of the Cognitive Sciences*, MIT Press. Cambridge. cxi-cxxxii. <https://doi.org/10.7551/mitpress/4660.003.0009>
- Tomasello, Michael 2010. *Origins of Human Communication*. MIT Press. Boston.

Geoff Vaughan

PhD Student, Full-time Lecturer
Eszterházy Károly Catholic University
Doctoral School of Education
E-mail: geoffvaughan1@gmail.com
<https://orcid.org/0000-0002-9614-1195>

Rita Szűts-Novák

Educational Researcher, Assistant Professor
Eszterházy Károly Catholic University
Institute of Digital Technology
E-mail: szuts-novak.rita@uni-eszterhazy.hu
<https://orcid.org/0009-0007-5199-047X>