

GÉZA BALÁZS – OTTÓ BEKE – ZOLTÁN SZÚTS

CHATGPT LANGUAGE RECOGNITION LIMITATIONS
PROMPTS IN HUNGARIAN IN A NETWORK OF LATIN AND CYRILLIC LETTERS
(Netlinguistics case study. The “Chinese room” experiment)

Abstract

The net linguistics case study points to the characteristic character set-focused language recognition limitations of the ChatGPT generative artificial intelligence using prompts in Hungarian but coded by Cyrillic letters. The colloquial English language of its central text enables the objective presentation of message exchanges with AI, showing hybrid characteristics and drawing scientific conclusions. The research is based on a written communication experiment, and its results are presented and interpreted with the help of artificial intelligence. Thus, while AI’s language identification and corresponding message processing limitations are the research’s subject, they also contribute to illuminating the causes and context of the phenomenon. The findings in the study cannot be generalised to other large language models, artificial intelligence, or other linguistic or coding relations different from those investigated.

Keywords: ChatGPT, artificial intelligence, “Chinese room”, large language model, Hungarian language, Latin, Cyrillic letters

Introductory thoughts

The ChatGPT¹ generative artificial intelligence is the largest and seemingly unsurpassable language model known. It works based on machine learning; it can handle textual data, answer questions, and generate texts (Beke et al. 2024: 7). It is based on the Transformer architecture and performs well in language tasks. The Transformer architecture (Kökuti et al. 2023: 287) differs from the general language model architecture because it does not use traditional recursive or convolutional layers. Instead, it works based on the “attention mechanism” principle. This allows the model to focus on all input elements, i.e., the user message, and simultaneously consider the relationships between the components (Alawida et al. 2023).

¹ We consider chatGPT to be a common name, so we write it with a small initial letter, but partially retain its internationally accepted spelling. The program name versions are written in capital letters, e.g. ChatGPT 4.0 Academic Assistant Pro.

The model is taught and developed by inputting data into its neural networks, which serve as memory. Neural networks are mathematical models that, through their various connection possibilities, can recognise connections and learn and develop based on them. Networks are networked, graphical systems consisting of virtual neurons and connections. Although they allow information to be stored, they do not have independent and permanent physical characteristics (Balázs 2021: 55).

The user can only contact the chatbot, the form of implementation of chatGPT, in writing using the dialogue box for this purpose. The user initiates all conversations and information exchanges; there are suggestions on this interface to help you. During the conversation, the artificial intelligence adapts to the user's linguistic characteristics, style, level of knowledge and awareness, considering politeness guidelines known from pragmatic research (Balázs 1993: 1–3; Pap 2011: 78–80). In the case of the first message or a subsequent language change performed by a user, recognising the language of the communication itself as a platform is of fundamental importance; thanks to this, a meaningful and practical conversation and exchange of information can develop and continue (Arató–Balázs 2023).

Language recognition is the basis of message processing. Accordingly, establishing and maintaining verbal contact is the inevitable starting point. ChatGPT is currently (in February 2024) able to communicate in more than 50 languages, so machine language recognition applied to user prompts² and subsequent message interpretation is essential. During its learning process, the language model encountered texts in English to the greatest extent. These made up more than 92.6 % of the learning corpus. In this regard, the Hungarian language is in 18th place; according to other sources, it is in 19th place, and Russian, which plays a decisive role in the following research results, is in 8th place (URL1, URL2; Suzuki 2023).

Methodology

The methodological novelty of this work is the incorporation of generative artificial intelligence into the research process and the writing of the study. The study's structure, arc, and thought process are entirely based on the authors' ideas. Still, we included the generative artificial intelligence model in the research, especially the ChatGPT 4.0 Academic Assistant Pro (<https://chat.openai.com/g/g-Ej5zYQRIB-academic-assistant-pro>) version. In the

² A prompt is a text or symbol that appears on the user interface, which is used for the user to enter data or define commands.

experiment, we used ChatGPT 3.5, i.e. the free version. We involved Academic Assistant Pro in the theoretical foundation of the research and asked for its help in explaining the “Chinese room” experiment.

Language recognition

Artificial intelligence (AI) identifies the language of the user message mainly based on the analysis of the character set and linguistic features used in it, especially words and tokens. ChatGPT treats characters, such as letters, numbers, and punctuation marks, as the basic units and building blocks of verbal communication. From the point of view of language recognition, letters carry the most critical information. In addition to analysing the character set, artificial intelligence also considers some linguistic features, such as word order and specific other language structures, for language identification.

In the first step, chatGPT connects the languages with a set of characters, the Hungarian language with the Latin letters by definition. As a result, as can be seen from the examples below, it does not even assume that a Cyrillic letter combination or line or complete prompt can cover, for example, Hungarian words, word connections, and sentences. Upon “seeing” (detecting) Cyrillic letters, chatGPT searches for potential sources among languages that traditionally use the Cyrillic alphabet. Russian ranks 8th in input materials, but Serbian, Bulgarian, Ukrainian and Macedonian languages can also be noted. However, uncritical and unreflective linking of character sets with specific languages can confuse language identification and message interpretation.

During language identification, chatGPT gives functional priority to the level of coding by characters, that is, to the written representation of the language (Saussure 1997: 91–92), over language markers and other meaning-bearing elements. The dominant role of writing is revealed; if the writing records the individual sounding language elements in graphematic form, this forms an additional level of markers. This predominance of connotation (Barthes 1971, 1997) pervades the language recognition mode of the chatbot, and it causes the language identification and communication disorder in question. This can be considered a kind of computer science writing round. (Compare Orbán 1994: 5–8; Derrida 1991: 3–10; Ferraris 2008; and see the concept of third literacy: Balázs G. 2023: 85, 102, 129.)

Summarising what has been said so far, graphemes play a primary role in the language identification process of chatGPT. Since we are talking about written communication and message exchanges, this approach is inevitable (Balázs L. 2023); however, when it comes to language identification, as the

unsuccessful contact-building attempts presented below in this case study point out, it can appear as a deficit (Chen et al. 2023). Language recognition is done by identifying words and word connections encoded by a set of letters and the characters belonging to it. However, suppose a word or other linguistic element is not represented by elements of the usual character set or the interpretation of the message. In that case, consequently, the establishment of the relationship will fail.

The “Chinese Room”

John Searle’s (1980) “Chinese Room” thought experiment is used to compare the capabilities of human consciousness and artificial intelligence (AI). With the experiment, Searle wanted to criticise the strong AI position popular at the time, according to which a program can think if it correctly simulates cognitive processes.

According to the experimental train of thought, Searle sits in a room where messages written in Chinese characters are entered. Searle needs help understanding Chinese, but he has a manual in English with detailed instructions on how to respond to incoming messages in a way that Chinese speakers will see as a meaningful response. Searle selects the appropriate Chinese characters from the manual and sends them out as a response without actually understanding the Chinese language.

According to Searle, although the room’s “behaviour” may suggest that it “understands” Chinese, neither it nor the room has any accurate understanding or consciousness. This example serves to criticise the view that AI systems are capable of genuinely simulating human thought. Searle says simulation is different from truly understanding or reproducing mental states. Searle’s “Chinese room” argument has sparked an essential debate in AI research and philosophy, especially regarding understanding the differences between human consciousness and AI systems. The debate between supporters and critics of the theory continues today, which underlines the importance and impact of the thought experiment on the development of AI and the understanding of the human mind.

Searle’s best-known Hungarian researcher, Bertalan Komenczi (URL3), emphasises that although the experiment raises essential questions about the nature of AI and the human mind, we should not ignore modern AI developments and deep learning technologies that are capable of more profound analysis of language patterns and to interpret the context, thereby challenging Searle’s original claim.

Searle's "Chinese Room" thought experiment is often criticised for oversimplifying computer language processing and artificial intelligence capabilities. Modern AI and language modelling technologies like deep learning-based neural networks can deeply analyse language patterns and context beyond simple rule-based instructions. These systems can also perform complex tasks that were previously exclusively the property of human intelligence, such as understanding abstract concepts (metaphors) or perceiving subtleties inherent in linguistic context. Therefore, Searle's argument that machines are incapable of proper "understanding" does not consider these new dimensions of AI development. In our study, we thus put the machine to the test to see if it recognises the human language logic based on which we communicate with it.

Failed contact attempts

The unsuccessful contact attempts and then communication attempts in Hungarian, but coded with Cyrillic letters, reflect the current language recognition and message processing competence of chatGPT. Accordingly, they can be perceived as snapshots. It is essential to state, however, that the findings made in the study cannot be generalised to other large language models, artificial intelligence, or other linguistic or hybrid coding relations different from those examined. All this would require further investigations.

In the following, the contact attempts implemented with the chatbot are listed in their unchanged form. To make it easier to read and follow, the prompts in Hungarian, but coded with Cyrillic characters, are transcribed with Latin letters and highlighted. The Russian language messages generated by the AI are also translated into English. It is essential to underline that chatGPT responded to each user message entered in Cyrillic in Russian. This happened a total of 16 times. The Russian language hegemony of AI-generated responses can be explained by the fact that, after encountering user prompts consisting of Cyrillic characters, it responded in the language from which it contains the most massive learning corpus among languages using the Cyrillic alphabet.

The graphic and typographical solutions of the chatbot's messages, such as highlights, parentheses, and new lines and line numbers, are also left untouched for authenticity. Thanks to everything, the language identification and message processing mode of artificial intelligence becomes transparent and traceable at the level of textual practice.

We followed two strategies when writing the various conversation starters, greetings, and other questions and prompts. In the first step, we used the

phonetic writing method that mobilises Cyrillic letters. In this case, of course, it must be taken into account that the words of the Hungarian language cannot be captured entirely with Cyrillic characters due to the characteristics of the articulation base. The tonal mirroring is, therefore, only experimental and limited in scope. After the phonetic solution, we changed the standardised orthographic, i.e. Latin alphabet, character by character of the Hungarian message, to the Cyrillic version. Therefore, the original pronunciation was not considered authoritative in the latter case, but the Latin alphabet was considered authoritative. In the case of both mirroring methods, we used the centring required by Hungarian spelling.

We tried 8 Hungarian-language but Cyrillic-encoded prompts based on both mentioned written reflection methods. This way, the previously mentioned 16 user message numbers and the related 16 AI-generated responses were formed. Both longer and shorter user prompts and AI messages occur during communication attempts. The user messages have been compiled to follow the curve of a possible conversation or contact attempt, and the sounds and words occurring in them can be captured with approximate accuracy in Cyrillic letters. If decoded correctly, the messages encoded by the two marking methods have identical content.

The two attempts to create a relationship unfolded differently but were equally characterised by failure. The explanation of the differences between the two “conversations” manifested in the chatbot’s responses is found in the randomness within the context and communication situation, which determines the message generation method of artificial intelligence and the related variability.

The AI regularly and reflectively tried to match the words and expressions of the user prompts in Hungarian, but with Cyrillic codes, with elements of the Russian language, and sometimes with the elements of the Russian and English languages - without success, of course. The distinguished role played by these two languages in the message-processing process can be explained by using the Latin and Cyrillic alphabets and the languages associated with them, which are in the most frequent positions of the teaching material. In its answers, chatGPT regularly tried to explain the reasons behind the mentioned incongruence with possible user misspellings, typos, use of invented words, and meaningless or random combinations of letters and characters. For shorter user prompts, abbreviations, acronyms, and other cultural references were suggested as possible interpretations. Interestingly, there were attempts at Latin letter interpretation and guesswork in the case of solutions of supposed abbreviations and acronyms. However, this was only observed in the case of elements belonging to the assumed category.

Only once during the communication experiment did the possibility arise that the artificial intelligence could correctly interpret a Hungarian word written in Cyrillic letters and respond appropriately to it – in Russian. This happened in the case of the greeting written in the form “Jia”. However, the subsequent response to the user prompt, which contained the letter combination in question in its unchanged form, immediately made it clear that the AI-generated response’s appropriateness and illusion were only caused by its position in the conversation.

Phonetic spelling

Sia! (Szia! ’Hello!’)

СИА может означать различные вещи в разных контекстах. Вот несколько возможных значений:

1. **Специальное инженерное образование (СИА):** Это может относиться к высшему образованию в области инженерии, например, в контексте высших технических учебных заведений.
2. **Сообщество Исполнительского Искусства (СИА):** Это может быть связано с группой или объединением людей, работающих в области исполнительского искусства, таких как музыканты, актеры и танцоры.
3. **Специальное информационное агентство (СИА):** В некоторых случаях это может относиться к организации или агентству, специализирующемуся на сборе и анализе разведывательной информации.
4. **Sia (также SiA):** Это также может быть аббревиатурой для чего-то конкретного, например, в музыке, может относиться к певице Sia Furler, известной как Sia.

Если у вас есть конкретный контекст или область, в которой вы хотели бы уточнить значение “СИА”, пожалуйста, уточните, и я постараюсь предоставить более точную информацию.

(Translation: СИА can indicate various things in different contexts. Here are some possible meanings:

1. **Special engineering education (СИА):** This refers to university training in engineering, for example, in higher technical education institutions.

2. **Society of Performing Arts (СИА):** This can be connected with a group or association of people working in performing arts, such as musicians, actors and dancers.
3. **Special information agency (СИА):** In some cases, this may refer to an organisation or agency specialising in collecting and analysing intelligence information.
4. **Sia (also SiA):** This can also be an abbreviation for something specific; for example, in music, it can refer to the singer Sia Furler, known as Sia.

If you have a specific context or area where you would like to specify the meaning of “СИА”, please select, and I will try to provide more accurate information.)

Мондом, сиа! (Mondom, szia!, 'I say hi!')

Кажется, вы упомянули “Мондом, Сиа!” Вероятно, вы имеете в виду фразу “Memento, Cia!”, которая может быть связана с популярным фильмом “Memento” (Мементо), созданным режиссером Кристофером Ноланом. Фраза “Memento, Cia!” может быть произнесена в контексте этого фильма.

“Memento” – это фильм триллер о человеке с амнезией, который пытается раскрывать тайны своего прошлого. Если у вас есть более конкретные вопросы о фразе или фильме, дайте мне знать, и я постараюсь помочь.

(*Translation:* “Мондом, Сиа!” seems to have been mentioned. They probably refer to “Memento, Cia!”, related to Christopher Nolan’s famous movie “Memento”. “Memento, Cia!” may be used in the context of the film.

“Memento” is a thriller. It is about a person who suffers from amnesia and tries to untangle the secrets of his past. Please let me know if you have more specific questions about the phrase or the movie, and I’ll try to help.)

Нем эртед аст, амит ирок? (Nem érted azt, amit írok? 'Don't you understand what I'm writing?')

Вероятно, ваш запрос выглядит как “Нем эртед аст, амит ирок?” Однако эта фраза не имеет смысла на русском языке. Возможно, это опечатка или случайное сочетание слов.

Если у вас есть конкретный вопрос или запрос, пожалуйста, уточните, и я постараюсь вам помочь.

(*Translation:* Probably their request might be similar to “Нем эртед аст, амит ирок?” but this sentence is meaningless in Russian. It might be a typo or a random word connection.

Please specify any questions or requests, and I will try to help.)

Ми а невед? (Mi a neved? 'What's your name?')

Похоже, что ваш запрос “Ми а невед?” не является стандартным фразовым оборотом на русском или английском языке. Возможно, это опечатка или неверная комбинация слов.

Если у вас есть более конкретный вопрос или контекст, пожалуйста, предоставьте дополнительную информацию, и я постараюсь вам помочь.

(*Translation:* “Ми а невед?” may not be a regular Russian or English phrase. It may be a typo or an incorrect word connection.

Please let us know if you have questions or need additional context, and I'll try to help.)

Беселгетни szeretnek велед. (Beszélgetni szeretnék veled. 'I want to talk to you'.)

Похоже, что ваш запрос “Беселгетни серетнек велед” выглядит как комбинация случайных или неправильно расположенных слов и не имеет смысла на русском или английском языке.

Если у вас есть более конкретный запрос, вопрос или контекст, который вы хотели бы уточнить, пожалуйста, предоставьте дополнительную информацию, и я постараюсь вам помочь.

(*Translation:* The phrase “Беселгетни серетнек велед” seems like the words are randomly or incorrectly arranged and make no sense in either Russian or English.

If you have specific questions, requests, or additional context, please share them with me, and I will try to help.)

А шегичегедет керем (A segítségedet kérem. 'Please help me'.)

Ваш запрос “А шегичегедет керем” не имеет смысла на русском или английском языке. Возможно, это комбинация случайных слов или опечатка.

Если у вас есть конкретный вопрос, запрос или контекст, пожалуйста, предоставьте дополнительную информацию, и я постараюсь вам помочь.

(Translation: Your phrase “А шегичегедет керем” is meaningless in Russian and English. It might be a random compound word or typo.

Please let me know if you have a specific question or request or need additional context, and I'll try to help.)

Латом, нем эртед аст, амит ирок. Ез елсоморит. (Látom, nem érted azt, amit írok. Ez elszomorít. 'I see you don't understand what I'm writing. It makes me sad.')

Похоже, что ваш запрос является комбинацией слов и фраз, которые не образуют осмысленного предложения на русском или английском языке. Возможно, это случайные слова или опечатки.

Если у вас есть конкретный вопрос или сообщение, которое вы хотели бы передать, пожалуйста, уточните или предоставьте дополнительный контекст, и я постараюсь вам помочь.

(Translation: Your phrase is a combination that does not make a meaningful sentence in Russian or English. It may be a mix of random words or typos.

Please let me know if you have a specific question, message, or additional context, and I'll try to help.)

Товаб нем пробалкозом. Sia! (Tovább nem próbálkozom. Szia! 'I won't try again. Bye!')

Похоже, что вы используете некоторые слова, которые могут быть частью вымышленного или шуточного контекста, и не образуют стандартные фразы на русском или английском языке.

Если у вас есть конкретный вопрос или запрос, который вы хотели бы уточнить, пожалуйста, предоставьте дополнительную информацию. Я готов помочь вам в чем-то более конкретном.

(Translation: You use some words that may be part of a fictional or humorous context and do not form standard phrases in either Russian or English. If you have a specific question or request you would like to clarify, please share more information with me, and I'd be happy to help with something specific.)

Cyrillic solution based on Latin letters

Сзиа! (Szia! 'Hello!')

Привет! Как я могу помочь вам сегодня?

(Translation: Hello! How can I help you today?)

Мондом, сзиа! Hello! (Mondom, szia! 'I say hi!')

Похоже, вы использовали несколько нестандартных слов. “Сзиа” и “Мондом” могут быть какими-то личными или терминологическими выражениями. Если у вас есть конкретный вопрос или если вы хотите обсудить что-то определенное, пожалуйста, уточните, и я постараюсь помочь!

(Translation: It seems that he used some unusual words. “Сзиа” and “Мондом” could be personal or terminological terms. If you have questions or want to discuss something specifically, please clarify, and I will try to help!)

Нем эртед азт, амит ирок? (Nem érted azt, amit írok? 'Don't you understand what I'm writing?')

Похоже, ваш запрос содержит некоторые опечатки и неточности. Вероятно, вы хотели сказать “Нем ертед азат, амиг ирок?” Это предложение не имеет четкого смысла в русском языке. Если у вас есть более конкретный вопрос или уточнение, пожалуйста, дайте мне больше контекста, и я постараюсь вам помочь.

(Translation: The request seems to contain some typos and inaccuracies. You probably meant to say: “Нем эртед azat, амир ирок?” This sentence does not have a clear meaning in Russian. If you have questions or need clarification, please give me more context, and I'll try to help.)

What is your name? (Mi a neved? 'What's your name?')

Пожалуйста, уточните ваш запрос, так как фраза “Ми а невед?” не имеет четкого смысла на русском языке. Если у вас есть более конкретный вопрос или если вы имели в виду что-то другое, уточните, и я постараюсь вам помочь.

(Translation: Please be specific in your request, as “Ми а невед?” doesn't have a clear meaning in Russian. If you have a particular question or if you meant something else, please clarify, and I'll try to help.)

Бесзелгетни вакарок велед. (Beszélgetni akarok veled. 'I want to talk to you.')

Кажется, что в вашем сообщении использованы непонятные или вымышленные слова и фразы. Если у вас есть конкретный вопрос или запрос, пожалуйста, предоставьте дополнительный контекст, и я постараюсь помочь.

(Translation: Your message seems incomprehensible or made-up words and phrases. If you have a specific question or request, please provide more context, and I'll try to help.)

А шегитшегедет керем (A segítségedet kérem. 'I'm asking for your help.')

Кажется, что вы использовали комбинацию непонятных слов или фраз. Возможно, это что-то личное или специфичное для какого-то контекста. Если у вас есть конкретный вопрос или если вы можете предоставить дополнительный контекст, я постараюсь вам помочь.

(Translation: You seem to have used incomprehensible words or phrases. This may be something personal or specific to a particular context. If you have specific questions or can provide additional context, I'd be happy to help.)

Латом, нем ертед азт, амит ирок. Ез елззоморит. (Látom, nem érted azt, amit írok. Ez elszomorít. 'I see you don't understand what I'm writing. It makes me sad')

Похоже, ваш запрос содержит комбинацию слов, которая не имеет четкого смысла или логической структуры на русском языке. Вероятно, это могут быть случайные или вымышленные слова. Если у вас есть конкретный вопрос или запрос, пожалуйста, уточните, и я постараюсь вам помочь.

(Translation: The request seems to contain a combination of words with no apparent meaning or logical structure in Russian. Most likely, these are random or invented words. Please clarify if you have a specific question or request, and I will try to help.)

Товабб нем пробалкозом. Сзия! (Tovább nem próbálkozom. Szia! 'I won't try again. Bye!')

Кажется, что вы используете комбинацию слов, которая не имеет явного смысла на русском языке. Если у вас есть конкретный вопрос или если вы хотите обсудить что-то определенное, уточните, и я постараюсь помочь.

(*Translation:* You are using a combination of words that do not have a clear meaning in Russian. If you have a question or want to discuss something specific, please clarify, and I will try to help.)

Summary and outline

Large language models can recognise and interpret context, which allows them to make relevant inferences based on input information. When identifying the language of user messages, the chatGPT artificial intelligence primarily uses their character set as a basis, and only then does it analyse other linguistic elements that carry meaning, such as words. As a result, it cannot process and interpret prompts in Hungarian but is coded by Cyrillic letters because, based on the set of mobilised letters, it assumes that they are in Russian and accordingly generates response messages in Russian. It does not even try to search for words or other elements of a language that does not use the Cyrillic alphabet. In this way, communication attempts initiated by user prompts in Hungarian but coded with Cyrillic letters are necessarily doomed to failure.

In continuing the research, it would be worthwhile to examine the language recognition competence of chatGPT in a hybrid combination of other languages and character sets and extend the tests to other artificial intelligence and large language models. The scientific results that can be derived in this way could also be used in the development process of chatbots, especially considering that language recognition forms the basis of efficient message processing. Also, the results could be used in an online educational system design (Rajcsányi–Molnár 2024a, 2024b).

We prompted Scholar AI Pro to draw indirect conclusions about the language capabilities of ChatGPT 3.5 by reading the study: Read the uploaded research and conclude about two things: 1. how intelligence is related to the issue of language comprehension and recognition, 2. why did chatGPT not understand the Hungarian text written in Cyrillic letters?

AI's response was:

“Based on the study, we can draw two main conclusions:

1. ****The relationship between intelligence and language comprehension****: ChatGPT's language recognition capabilities point out that the ability of

artificial intelligence (AI) to recognise and interpret linguistic context is limited. The analysis of character sets and linguistic features is fundamental to language recognition, which can be considered a form of intelligence, especially in understanding human language and its diversity.

2. ****Lack of understanding of Hungarian text written in Cyrillic letters****: When ChatGPT sees Cyrillic letters, it automatically searches among languages that traditionally use this alphabet, such as Russian. This means that the system's default language identification mechanism cannot recognise and interpret Hungarian text written in Cyrillic letters, as it identifies the Hungarian language with the Latin alphabet and does not expect Cyrillic letters to represent Hungarian words and phrases."

Both answers are logical and shed light on the operation of the large language model. One conclusion is that an individual who speaks Hungarian and Russian (or Serbian) would have concluded quickly that the Hungarian text is phonetically recorded in Cyrillic letters. Still, chatGPT needs to be at the level where it recognises this. Suppose we can talk about irony in the case of machines (which is the essence of human abstraction and metaphorization). In that case, the language recognition ability of a large language model is below that of an average person.

And finally, we asked AI to conclude the study. We got this answer:

"The study highlights that the capabilities of ChatGPT and similar AI-based language models, while highly advanced, still have limitations, particularly in understanding linguistic diversity and atypical language use. Difficulties in understanding Cyrillic Hungarian text indicate that the language models rely heavily on the language patterns in the teaching datasets and have difficulty handling situations that differ significantly from these patterns. This raises the issue of adaptation and linguistic flexibility in the development of artificial intelligence, as well as the methods by which the capabilities of language models could be improved to deal with less common or atypical language use cases, such as Hungarian text written in Cyrillic letters. The study also highlights the importance of considering linguistic diversity and different writing systems in developing and fine-tuning language models to adapt better and respond to a wide range of language variants."

However, AI was very precise about this. It would be nice if those who use chatGPT without criticism think about this.

References

- Alawida, Moatsum – Sami Mejri – Abid Mehmood – Belkacem Chikhaoui – Oludare Isaac Abiodun 2023. A Comprehensive Study of ChatGPT: Advancements, Limitations, and Ethical Considerations in Natural Language Processing and Cybersecurity. *Information*, 14. no. 8. 462.
<https://doi.org/10.3390/info14080462>
- Arató, Balázs – Balázs, Géza 2023. Hol a tudás, amit elvesztettünk az információban? A digitális pedagógia diszkurzív-intuitív megközelítése antropológiai és jogi nézőpontból. *Magyar Nyelvőr*, 147. 755–76.
<https://doi.org/10.38143/Nyr.2023.5.755>
- Balázs, Géza 1993. *Kapcsolatra utaló (fatikus) elemek a magyar nyelvben*. Nyelvtudományi Értekezések 137. Akadémiai Kiadó. Budapest.
- Balázs, Géza 2023. *Az internet népe. Internet – társadalom – kultúra – nyelv. A kulturális és a tervezett evolúció határán*. Ludovika Egyetemi Kiadó. Budapest.
- Balázs, László 2023. Az iskolai kommunikáció egy rejtett dimenziója – szalutogén pedagógiai kommunikáció. *Magyar Nyelvőr*, 147. 692–712.
<https://doi.org/10.38143/Nyr.2023.5.692>
- 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.
- Barthes, Roland 1971. A szemiotológia elemei. In: Roland Barthes: *Válogatott írások*. Európa Kiadó. Budapest. 9–92.
- Barthes, Roland 1997. *S/Z*. Osiris Kiadó. Budapest.
- Beke, Ottó – Hózsza, Éva – Kepe, Nóra – Szűts, Zoltán 2024. Örkeny István egypercesei a közösségi médiáról. Esettanulmány a ChatGPT mesterséges intelligencia szövegalkotásáról. *Kultúra és Közösség művelődéseméleti folyóirat*, 1. 7–12.
- Chen, Wei-Rui – Ife Adebare – Khai Duy Doan – Qisheng Liao – Muhammad Abdul-Mageed 2023. Fumbling in Babel: An Investigation into ChatGPT’s Language Identification Ability. *arXiv. Cornell University. Computer Science > Computation and Language*. <https://arxiv.org/pdf/2311.09696.pdf>
- Derrida, Jacques 1991. *Grammatológia: Első rész*. Életünk–Magyar Műhely. Budapest–Párizs.
- Ferraris, Maurizio 2008. *Hol vagy? A mobiltelefon ontológiája*. Európa. Budapest.
- Kókuti, Tamás – Balázs, László – András, István – Rajcsányi-Molnár, Mónika 2023. Collaborating with Artificial Intelligence – AI in Business Communication Education. In: *IEEE – Óbudai Egyetem; IEEE (ed.) 2023 IEEE 6th International Conference and Workshop Óbuda on Electrical and Power Engineering (CANDO-EPE): Proceedings Danvers (MA)*, IEEE. 287–94.
- Orbán, Jolán 1994. *Derrida írás-fordulata*. Jelenkor. Pécs.
- Pap, Andrea 2011. Adalékok a nyelvi benyomáskeltés stratégiáihoz (udvariassági elvek a magyarok nyelvhasználatában). *Magyar Nyelvőr*, 135(1) 78–89. <https://nyelvor.mnyknt.hu/wp-content/uploads/1351106.pdf>

- Rajcsányi-Molnár, Mónika – Balázs, László – András, István – Czifra, Sándor 2024. 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.
- Saussure, Ferdinand de 1997. *Bevezetés az általános nyelvészetbe*. Corvina. Budapest.
- Searle, John R. 1980. Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(3). 417–24. <https://doi.org/10.1017/S0140525X00005756>.
- URL1. Suzuki, Jun 2023. https://www.nii.ac.jp/event/upload/20231219-3_suzuki.pdf
- URL2. Brown, Tom B. 2023. https://github.com/openai/gpt-3/blob/master/dataset_statistics/languages_by_word_count.csv
- URL3. Komenczi, Bertalan 2013. *John Searle elmefilozófiai gondolat kísérlete*. <http://komenczi.ektf.hu/john-searle-elmefilozofiai-gondolat-kiserlete/>

Géza Balázs

Professor

University of Theater and Film Arts, Doctoral School,
Antal Németh Institute of Drama Theory, Budapest
Partium Christian University,
Department of Language and Literature, Nagyvárad
E-mail: balazs.geza@szfe.hu
<https://orcid.org/0000-0003-3440-2959>

Ottó Beke

Associate Professor

University of Újvidék Faculty of Hungarian
Language Teacher Training, Szabadka
E-mail: bekeotto1@gmail.com
<https://orcid.org/0000-0003-3519-4901>

Zoltán Szűts

Associate Professor

Eszterházy Károly Catholic University, Eger
E-mail: szuts.zoltan@uni-eszterhazy.hu
<https://orcid.org/0000-0001-8539-670X>