THE ROLE OF DIGITAL CORPORA, ARTIFICIAL INTELLIGENCE, AND LINGUISTIC MODELING IN CONTEMPORARY LINGUISTIC RESEARCH

Umarova Marhabo Erkinovna
2326, First English Faculty UzSWLU
marhaboumarowaa@gmail.com

Annotation. This paper examines the function of digital corpora, artificial intelligence, and linguistic modeling in contemporary linguistic research. It is founded on a qualitative analysis of scholarly literature in corpus linguistics and computational linguistics. The results show that digital corpora give real language data for analysis and that artificial intelligence helps people learn languages by using adaptive systems and automated feedback. Linguistic modeling helps put language structures into a form that computers can understand. The study also points out problems, like too much reliance on technology and less interaction between people. In general, people see these tools as helpful rather than replacing traditional methods.

Keywords: digital corpora, artificial intelligence, linguistic modeling, corpus linguistics, natural language processing, adaptive learning

Аннотация. В статье рассматриваются цифровые корпусы, искусственный интеллект и лингвистическое моделирование в современной лингвистике. Работа основана на анализе научной литературы. Результаты показывают, что цифровые корпусы предоставляют реальные языковые данные, искусственный интеллект применяется в обучении языкам, а моделирование помогает анализировать структуру языка. Также отмечаются ограничения технологий и необходимость участия человека.

Ключевые слова: цифровые корпусы, искусственный интеллект, лингвистическое моделирование, корпусная лингвистика, компьютерная лингвистика, обработка естественного языка

Annotatsiya. Ushbu maqolada raqamli korpuslar, sun’iy intellekt va lingvistik modellashtirishning zamonaviy tilshunoslikdagi o‘rni ko‘rib chiqiladi. Tadqiqot ilmiy adabiyotlar tahliliga asoslangan. Natijalar shuni ko‘rsatadiki, raqamli korpuslar til haqida haqiqiy ma’lumotlarini beradi, sun’iy intellekt til o‘qitishda qo‘llaniladi, modellashtirish esa til tuzilmasini tahlil qiladi. Texnologiyalarning cheklovlari va inson omilining ahamiyati ham ta’kidlanadi.

Kalit so‘zlar: raqamli korpuslar, sun’iy intellekt, lingvistik modellashtirish, korpus lingvistikasi, kompyuter lingvistikasi, tabiiy tilni qayta ishlash

Introduction

The study of linguistics has changed a lot in the last few decades. In the past, a lot of linguistic research was based on gut feelings, theoretical examples, or small sets of data. Researchers now work with a lot of real language data and use digital tools to help them analyze it. This change has made linguistic research more based on facts and, in many cases, more trustworthy.

The utilization of digital corpora[1] is a significant advancement in this field.These are big groups of real texts that show how people actually use language in real life. Researchers can now look at real patterns of language use, such as word frequency, collocations, and grammatical structures in different situations, instead of making up fake examples (Sinclair, 1991). This makes it easier to see small differences in meaning and use that you might not have noticed before. Artificial intelligence has also become more common in language

DOI: 10.5281/zenodo.20377759

research and language learning. AI tools can now read texts, give feedback, and even pretend to talk to users. For instance, language learners can talk to chatbots or get automatic corrections, which makes the learning process more interesting and easier to access.

In this case, linguistic modeling is also very important. It helps computers understand language by putting it in a structured and systematic way (Gries, 2015). This is especially important for things like speech recognition and machine translation, where you need to know how language works.

In general, these changes show that linguistics is no longer limited to old ways of doing things. It is becoming more based on data and more connected to technology. This paper aims to examine the utilization of digital corpora, artificial intelligence, and linguistic modeling in contemporary linguistic research, while evaluating their advantages and limitations.

Methods

This study is predicated on a qualitative analysis of academic literature rather than original empirical research. A variety of sources were utilized, encompassing books, peer-reviewed journal articles, and review studies pertinent to corpus linguistics, computational linguistics, and artificial intelligence in education.

The sources were chosen because they were important to the field and related to the topic. We focused on studies and works that are well-known and explain clearly how digital tools are used in linguistic research. For example, because they are frequently used in linguistic studies, materials discussing well-known corpora like the British National Corpus (BNC) and the Corpus of Contemporary American English (COCA) were included. Additionally, research on AI-based language learning tools was examined. These included studies on platforms for adaptive learning, chatbots, and automated writing assessment systems. These resources are especially pertinent to this study because they are frequently utilized in actual educational environments.

The goal was to compare results from various sources and find common patterns rather than concentrating on a single experiment. Instead of focusing on specific outcomes, this method allows for the observation of broad trends. It also helps to understand how digital corpora, artificial intelligence, and linguistic modeling are connected in practice (McCarthy, 2004).

Results

The examination of the chosen literature indicates that digital corpora, artificial intelligence, and linguistic modeling have emerged as significant components of contemporary linguistic research.

First and foremost, digital corpora give you access to big collections of real language data. Researchers can study how language is used in real situations instead of making up examples. Corpora, for example, let you look at how often certain words show up, which words tend to show up together, and how different types of texts use grammatical structures (Sinclair, 1991). This makes linguistic analysis more objective and based on evidence.

AI is also becoming more important, especially for learning languages. Chatbots, grammar checkers, and automated feedback systems are some of the tools that help students practice a language and get quick answers. This can make learning more interesting and adaptable because students can work at their own pace and focus on what they need (Zawacki-Richter et al., 2019). AI tools can also help teachers by cutting down on the time it takes to do everyday tasks like checking homework.

On the other hand, linguistic modeling is necessary to link language with technology. It lets computers understand language by putting it in a format they can read. This is

especially important for things like machine translation, speech recognition, and text generation (Gries, 2015).Computers would have a hard time "understanding" or working with human language if there weren't any linguistic models. The analysis also shows some problems, though. One of the biggest worries is that relying too much on technology could make it harder for people to interact with each other in real life. AI systems also don't always fully understand the context, which can cause them to make mistakes or use language that sounds strange.

Discussion

The findings demonstrate the interdependence of digital corpora, artificial intelligence, and linguistic modeling, all of which influence current linguistic research.

Each element makes a distinct contribution: linguistic modeling provides structural frameworks, artificial intelligence permits processing and interaction, and corpora supply empirical data. Both research and language instruction are improved by their integration. Human interpretation is still crucial at the same time. Large datasets can be processed by machines, but they are unable to fully capture contextual nuance, communicative intent, or cultural meaning (McCarthy, 2004). Therefore, rather than taking the place of conventional methods, technology should enhance them.

Future studies will probably concentrate on further integrating these tools while preserving a balance between human expertise and technological efficiency.

References

1. Baker, P. (2006). Using corpora in discourse analysis. Continuum.
2. Gries, S. T. (2015). Quantitative corpus linguistics with R. Routledge.
3. McCarthy, M. (2004). English corpus linguistics: An introduction. Cambridge University Press.
4. Sinclair, J. (1991). Corpus, concordance, collocation. Oxford University Press.
5. Zawacki-Richter, O., et al. (2019). Systematic review of artificial intelligence in higher education. International Journal of Educational Technology in Higher Education, 16(1), 1–26.