



MIRZO ULUG'BEK NOMIDAGI
O'ZBEKISTON MILLIY UNIVERSITETINING
JIZZAX FILIALI

**ZAMONAVIY INNOVATSION
TADQIQOTLARNING
DOLZARB MUAMMOLARI
VA RIVOJLANISH
TENDENSIYALARI:
YECHIMLAR VA ISTIQBOLLAR
RESPUBLIKA ILMIY-TEXNIK
ANJUMAN MATERIALLARI
TO'PLAMI**



**15-16-MAY
2026-YIL**



**Google
Scholar**



**O‘ZBEKISTON RESPUBLIKASI OLIY TA’LIM, FAN VA
INNOVATSIYALAR VAZIRLIGI**

**MIRZO ULUG‘BEK NOMIDAGI O‘ZBEKISTON MILLIY
UNIVERSITETINING JIZZAX FILIALI**

**ZAMONAVIY INNOVATSION TADQIQOTLARNING DOLZARB
MUAMMOLARI VA RIVOJLANISH TENDENSIYALARI: YECHIMLAR
VA ISTIQBOLLAR**

*mavzusidagi Respublika ilmiy-texnik anjuman materiallari to‘plami
(2026-yil 15-16-may)*

JIZZAX-2026

1. Вересаев В.В. «Сёстры» [Электронный ресурс]: Электронная библиотека RoyalLib.Com». - Электрон. библиотека — режим доступа: https://royallib.com/read/veresaev_v/sestri.html(дата обращения: 16.10.2025).
2. Степанов С.С. Язык внешности. Жесты, мимика, черты лица, почерк и одежда. – Москва : Эксмо-Пресс, 2001. – 416 с. – с. 363. – ISBN: 5-04-005684-2
3. Телия В. Н. Русская фразеология: Семантический, прагматический и лингвокультурологический аспекты. – М.: Языки русской культуры, 1996. – 288 с.
4. Толковый словарь русского языка. Т. 1. А — Кюрины / гл. ред. : Б. М. Волин, Д. Н. Ушаков ; сост. : В. В. Виноградов, Г. О. Винокур, Б. А. Ларин, С. И. Ожегов, Б. В. Томашевский, Д. Н. Ушаков ; под ред. Д. Н. Ушакова. – Москва : Гос. изд-во иностр. и нац. слов., 1935. – 828 с. – с.868
5. Толковый словарь русского языка. Т. 3. П – Ряшка / гл. ред. : Б. М. Волин, Д. Н. Ушаков ; сост. : В. В. Виноградов, Г. О. Винокур, Б. А. Ларин, С. И. Ожегов, Б. В. Томашевский, Д. Н. Ушаков ; под ред. Д. Н. Ушакова. – Москва : Гос. изд-во иностр. и нац. слов., 1939. – 714 с. – с.296

BRIDGING THE MORPHOLOGICAL GAP: COGNITIVE PROCESSING, READING EFFICIENCY, AND COMPREHENSION STRATEGIES IN HIGHLY AGGLUTINATIVE LANGUAGES

Jo‘rayev Muhammadrahimxon Murod o‘g‘li
Mirzo Ulug‘bek nomidagi O‘zbekiston Milliy universitetining Jizzax Filiali
mukhammadrahimkhonjuraev@gmail.com

One of the main factors for non-native speakers having problems in understanding texts written in agglutinative languages is the lack of knowledge of the morphological rules that are used in these languages. The syntactic structures of non-agglutinative languages, which mainly rely on the word order to express the meaning, are often very rigid and do not present any problem for non-native speakers. However, agglutinative languages present a lot of syntactic variation that is mainly solved by the morphology. The use of affixes (prefixes, root vowel changes, infixes and suffixes) to form new words that express a wide range of grammatical meanings, often within a single word, makes the words of agglutinative languages very complex and very long. Thus, the main problem that non-native speakers have when understanding texts written in agglutinative languages is the decoding of the words that are used in these texts. The number of words that a non-native speaker has to learn in order to be able to read and write a text is, therefore, very high, and this number can be even higher if the non-native speakers do not have enough practice in using the morphological patterns of the target language that are needed to form words that are used in the texts that they read.

Morphological complexity directly influences readability. Because of the complexity of the semantic-dense morphological forms, non-native readers apply many

different strategies in order to process and understand the text they read. However, when the word length of the semantic-dense morphological forms exceeds the processing capacity of the reader, there is an increase in the number of errors on the part of the reader and also in the number of miscommunications that take place as a result of the reader's inability to understand the text that he or she is reading.

In addition, many agglutinative languages have great syntactic diversity. This means that even a short sentence can be constructed in many different ways. The word order can be changed, depending on the clause type and even on the emphasis of the sentence. Moreover, one word can have different grammatical functions depending on its morphological structure. As a result, many sentences can be interpreted in different ways and only one of them can be correct. A non-native speaker of an agglutinative language will often not know which syntactic construction to apply and how to interpret a sentence. Thus, even he or she is familiar with the individual elements of a sentence, he or she may have problems in combining them in order to understand the whole message. This means that non-native speakers of an agglutinative language will often not be able to understand a sentence even if they know every single word of that sentence. In such cases, they would need more syntactic knowledge and experience than they already have in order to understand the intended message.

Readability models are statistical models that attempt to predict the difficulty of reading a piece of written text. In the case of agglutinative languages, it is also possible to incorporate information that pertains to the morphological structure of the words in the text. Traditionally, the Flesch-Kincaid readability tests have used the length of the sentences as well as the average number of letters and words per sentence in order to predict the readability of a piece of written text. Isolating languages, which generally do not make use of morphology, are typical cases in which these measures are effective. However, the case of agglutinative languages is clearly very different. Thus, in order to make effective use of readability models in agglutinative languages, it is necessary to incorporate information that pertains to the morphological structure of the words in a text. The most common method for doing this is through the use of the Morpheme Count, or the total number of individual morphemes in a given piece of written text. This measure is effective, because the total number of individual morphemes in a piece of written text serves as a good predictor of the overall amount of cognitive load that the reader will need to expend in order to read the text in question.

New methods of measuring the readability of a language have been developed in recent years. The majority of these methods take into account a number of aspects, such as the length of words and sentences, as well as the proportion of simple words within a text. When it comes to agglutinative languages, however, there are several specialized methods for calculating readability. The Morphological Readability Index (MRI) and the Agglutination Readability Model (ARM) are two examples of such metrics, which have been implemented in a number of studies in order to measure and analyze the various aspects of a given language's readability. As was demonstrated in a study conducted by Kh et al. (2020), the agglutination rate of a given language has a significant negative correlation with readability, which suggests that in general, agglutinative languages are more difficult to read than languages that employ non-agglutinative morphological structures.

The impact of linguistic indices extends beyond the area of readability to also include assessment of the syntactic diversity of agglutinative languages. A text written in an agglutinative language may be very complex due to the large number of different agglutinative formations, such as those that contain a rich case system and/or a large number of verb inflections. While this complexification of language can be seen as a positive feature of the language, which makes possible a great number of different expression of the same idea, it can also pose difficulties to the learners of the language. The complexity of the language can affect the scores of non-native speakers on standardized reading comprehension tests, when the morphosyntactic complexity of the text is measured by the amount of syntactic variation. Such variation can indeed be highly correlated with the scores of non-native speakers on reading comprehension tests.

Recent advances in computational linguistic research have greatly contributed to a better understanding of issues faced by agglutinative languages with complex morphological structures. The syntax of agglutinative languages is particularly difficult for computational processing (parsing, similarity detection, analysis), and has been tackled with the help of recent advances in NLP (Natural Language Processing) research and machine learning (ML). Modern computational models for processing agglutinative languages are able to employ sophisticated morphological analyzers (dis- and re-synthesizers), improved methods for text similarity detection, as well as advanced measures for calculating syntactic diversity of a text (Biloshchytska et al., 2025).

Morphological analyzers and synthesizers for agglutinative languages have also undergone enhancements lately. To a large extent, this is related to improving readability of such languages by breaking down words into their component morphemes, i.e. performing syntactic parsing. One of the recently developed tools of this type is a neural network for morphological parsing that is able to process a high number of unique patterns formed by a large number of morphemes. Furthermore, deep learning architectures have also been extended to decomposition of words into their constituent subwords (Biloshchytska et al., 2025).

Moreover, in computational linguistic context, the new methods for detection of text similarity for languages with complex word-formation are of growing interest. In particular, traditional similarity metrics fail when dealing with agglutinative languages. However, in recent years a number of novel methods for representation of texts in vector space have been proposed, such as models based on word embeddings that model semantic relationships within a vector space. The FastText model and GloVe model have recently been successfully adapted for the agglutinative languages and thus can be effectively used for analysis of such languages on the subword level.

In summary, computational approaches for analyzing agglutinative languages greatly depend on current research in the field of computational linguistics and on methods for processing and analyzing complex, highly-morphemic languages. Therefore, recent improvements in sophisticated analysis tools for the decomposition and re-synthesis of words into their individual morphemes, high-quality representations of complex linguistic structures, better methods for determining text similarity, sophisticated tools for determining syntactic diversity in languages and

various other properties greatly aid researchers in their study of relationships between the morphological structures of texts and such aspects of reading as difficulty and comprehension. The study has demonstrated how morphological structures in agglutinative languages affect various dimensions of text, including complexity, readability, syntactic diversity and reading difficulties experienced by both native and non-native speakers of these languages. The complex morphology of languages like Turkish, Finnish and Swahili increases the syntactic complexity of language and causes difficulties to speakers of languages with simple morphology in reading and writing these languages. Though our research established a connection between the morphology of a language and the complexity of corresponding texts, there are plenty of new research questions. Above all, the relation of a language's morphology to its speakers' reading comprehension as well as to their effective use of corresponding teaching strategies deserves to be scrutinized in greater depth. And since the range of morphological types and their corresponding realization in agglutinative languages is vast, we must by all means investigate how native and non-native speakers deal with corresponding languages as a function of the languages' typological characteristics, i.e. in terms of a variety of potential strategies.

Used literature

1. Adali, E., & Adamov, A. Z. (2016, October). Sentiment analysis for agglutinative languages. In 2016 IEEE 10th International Conference on Application of Information and Communication Technologies (AICT) (pp. 1-3). IEEE.
2. Baykara, B., & Güngör, T. (2022). Abstractive text summarization and new large-scale datasets for agglutinative languages Turkish and Hungarian. *Language Resources and Evaluation*, 56(3), 973-1007.
3. Bekarystankyzy, A., Mamyrbayev, O., & Anarbekova, T. (2024). Integrated end-to-end automatic speech recognition for languages for agglutinative languages. *ACM Transactions on Asian and Low-Resource Language Information Processing*, 23(6), 1-17.
4. Biloshchytska, S., Tleubayeva, A., Kuchanskyi, O., Biloshchytskyi, A., Andrashko, Y., Toxanov, S., ... & Sharipova, S. (2025). Text similarity detection in agglutinative languages: A case study of Kazakh using hybrid n-gram and semantic models. *Applied Sciences*, 15(12), 6707.
5. Gósy, M., & Abari, K. (2021). The interaction of word complexity and word duration in an agglutinative.
6. Kambarami, F., McLachlan, S., Bozic, B., Dube, K., & Chimhundu, H. (2021). Computational modeling of agglutinative languages: The challenge for Southern Bantu languages.
7. Kargin, T., Guldenoglu, B., Gengec, H., & Toker, M. (2023). The importance of morphological knowledge in the reading comprehension difficulties in a highly agglutinative language: Evidence from poor comprehenders. *Journal of Psycholinguistic Research*, 52(2), 653-673.
8. Kh, S. S., Barahoeva, N. M., Kh, K. Z., & Dudarova, L. M. (2020). SOCIAL DIFFERENTIATION OF LANGUAGES WITH THE ANALYSIS OF AGGLUTINATION INDICES. *Revista Turismo Estudos e Práticas-RTEP/GEPLAT/UERN*, (4), 1-7.

9. Ladányi, E., Kovács, Á. M., & Gervain, J. (2020). How 15-month-old infants process morphologically complex forms in an agglutinative language?. *Infancy*, 25(2), 190-204.

10. Murzin, F. A., Sagnayeva, S. K., Yerimbetova, A. S., & Sambetbayeva, M. A. (2016). Agglutinative languages with a link grammar. *Вестник Казахской академии транспорта и коммуникаций им. М. Тынышпаева*, (2), 61-61.

UNRAVELING COMPLEXITY: THE ROLE OF MORPHOLOGICAL STRUCTURE IN READABILITY AND COMPREHENSION CHALLENGES OF AGGLUTINATIVE LANGUAGES

Jo‘rayev Muhammadrahimxon Murod o‘g‘li

Mirzo Ulug‘bek nomidagi O‘zbekiston Milliy universitetining Jizzax Filiali
mukhammadrahimkhonjuraev@gmail.com

Agglutinative languages, by their morphological properties, form words by a systematic use of affixes to a root. Therefore, many complex words can be formed by prefixes and/or by suffixes and/or by an infix, and in these languages each affix has one fixed grammatical function. For example, the root word of many Urdu verbs is to speak or to say. A past tense is added to this root as a prefix; and in many cases, cases are added to the root as suffixes; and mood is indicated by yet another suffix. This structure of the words of an agglutinative language does not fit within the familiar framework of analysis which is suitable for isolating languages or for so-called fusional languages (Sands, 2009). Hence, an special analysis of the complexity of texts in such languages, or of so-called ‘readability’ of a given text, or of its ‘syntactic diversity’ is called for. It is noted here that for purposes of this article, agglutinative languages are Turkish, Finnish and Swahili, the language of Zanzibar and the Zairian Congo. Here, it should also be added that as is the case with any class of languages, there are variations within the class of agglutinative languages. That is, however, a matter that shall have to be ignored in the present article. It is obvious that from a study of the morphological features of a language and its texts, lots of interesting questions could be formulated concerning issues of readability of written texts and their syntactic diversity for native speakers of these languages as opposed to that of non-native speakers. For instance, one would expect to find very different levels of complexity when such non-native speakers have had quite different amounts of formal study of the language, or at any rate of the grammar of that language. Also, even within any one language community, there must clearly exist a great many very different answers to each and every one of the aforementioned questions, depending on as many variables as the range of circumstances of life. These variables could range from a native speaker’s or a non-native speaker’s experience of reading or listening to written or spoken text, within that language’s speaking community, to the kind of situation in which any given piece of writing is being read by that person (Suleimanov et al., 2024).