



MODELING LEXICAL TOOLS IN DETERMINING THE SYNONYMY OF SIMPLE SENTENCES IN SCIENTIFIC TEXTS IS A TOPICAL ISSUE

Abror Murtazaev

Independent researcher at the Tashkent State University of Uzbek

Language and Literature named after Alisher Navoi

Email: abror.murtazayev.89@inbox.ru

Тел.99)389-03-07

Article history:

Received: 23rd August 2021

Accepted: 17th September 2021

Published: 21st October 2021

Abstract:

The development of research in Uzbekistan, the emergence of original scientific texts requires the development of a certain linguistic control (to prevent duplication). For this, it is necessary to create software that determines the level of similarity of the content of scientific texts available in electronic form. Experts in a given field, such as painters or sculptors, determine whether a work of art is copied, and scientists in a given field determine whether a work of art is copied. The increase in the number of problems in this example is related to the posting of research results in higher education on the Internet. This situation has arisen today and requires a new level of struggle against illegal use of someone else's intellectual property in the world.

Keywords: Identity, similarity, difference, lexical means, speech, artificial intelligence, linguistics, molding, scientific text, synonymy, content, copyright, linguistic expertise, linguistic model.

In linguistics, we have noted that while the current situation with copying and its prevention, i.e. the legal basis for the protection of copyright [12], is sufficient, scientific approaches in this regard are far behind. In philosophy, the effect of quantitative change on qualitative change is considered to be a law, but in the thirty years of independence of the Republic of Uzbekistan, especially in linguistics, there are no solutions to current problems that can compete with world linguistics. Of course, the law is strong, but human needs are even stronger. As we mentioned at the beginning, the law can enforce what can be done, what cannot be done, and what is punishable when caught doing a prohibited act. However, it is unreasonable and impossible to expect an effective result if copyright infringement is not scientifically substantiated and is not divided into methods and tools of in-depth analysis. Today, even in developed countries, determining the original text of research is a serious problem. The fact that higher education institutions post scientific papers on the Internet and find such papers cheaply and easily using the Internet is exacerbating the problems in this area. Of course, as mentioned above, there are many commercial programs abroad that check the authenticity of scientific texts. However, the fact that there are more commercial offers against it, in particular, "smart synonymizers", as well as the relentless production of methods of cheating the program is a serious and topical issue. Using online anti-plagiarism software on the Internet, we have developed our own proposals and recommendations for the preparation of a linguistic base for anti-plagiarism software, which determines the level of semantic similarity of simple sentences in Uzbek scientific texts, and to determine its working methods.

Language molding serves to ensure that it is used automatically, adapting it to allow the simplest tasks to be performed by a human being, as well as tasks that require a lot of time and a large amount of memory, quickly and flawlessly. A linguistic model is an intellectual tool that is artificially developed by a linguist to replicate or imitate his or her behavior and the behavior of the original for linguistic purposes. There are many definitions of models in linguistics. Most often, this term is understood as follows:

- model - type, sample (language sample) of any text units (words, sentences);
- model - symbols, schemes for describing language objects (PR: scheme of the component model in syntax),
- model - a structure theory formalized with a defined metal language (PR: formal grammars). [15]

The main purpose of modeling in linguistics is to model for the artificially convenient use of a person's integral language ability.

The concept of linguistic model originated in structural linguistics, but was introduced into scientific circulation in the 60s and 70s.

In computer linguistics, the term linguistic module plays an important role today. This is due to the transfer of natural language to computer language, that is, the development of ways to process text through a computer system.

In this regard, linguistic programs of foreign languages have been developed and are being improved today. The linguistic module is an independent component of such linguistic software, that is, the part of the software that covers a particular linguistic process. [1] In fact, language theory is also derived from the very nature of the existing language, preparing its specific aspects for systematic use according to a certain order. That is, the theory originated from practice, and today the phenomenon of the return from theory to practice is much more active.

By downloading scientific works available on the Internet through anti-plagiarism programs available on the Internet (anti-plagiarism programs on the Internet compare them with the texts in their hidden and downloaded databases on the Internet), we identify the similarities in the texts in the following steps:

1. The stage of purification, in which the removal of excess or aspects and means that interfere with the verification of the originality of the text.

1.1. The filled parts are identified by taking a screenshot of the text. The reason is that in the 12 secrets of cheating the anti-plagiarism system in 2021 [17], this method was proposed as a recommendation (until 2016 it worked like putting various invisible characters between text, replacing letters with Cyrillic and Latin).

1.2. Clearing the text from spelling mistakes is also relevant to this content, as today's existing programs only check the accuracy, that is, the exact copy, so changing a single word in each sentence can give a hundred percent result. For example, in scientific texts available on the Internet, the same thing happened when we changed the letter x to h and the letter o to u.

1.3. Clearing links is important at this point in determining how many links are possible relative to the total text. (Doesn't read quotes as plagiarism either.)

1.4. Non-syntactic parts of speech:

1.4.1. simple and complex entries are given in parentheses;

1.4.2. the introductory words, for example, are indefinite, indefinite, indeed, indeed, probably, probably, approximately (List is attached), etc. ;

1.4.3. Verbs formed by the affixes -la and -lan are removed from them, taking into account that they form a doublet relationship. For example, the verb to do can be done with verbs that are made with the help of -la: to exchange-to exchange, to force-to force, to spend-to spend, to confirm-to confirm (the list is attached), and so on. Insignificant tools in this section are coded NM1, NM2, and the like;

1.5. insignificant suffixes: -dir (when it comes to words in a group of words other than the verb), in the formation of word combinations by adaptation our house is our house, your house is your house i.e. in -miz and ning-zero form, ning-ingiz;

1.6. the connectors that come at the beginning of the sentence are also removed.

1.7. Clearing from the words of Sheva. Indeed, scientific texts must be expressed in words specific to literary language.

In modern linguistics, the meaning of the term "model" has largely been previously covered by the term "theory" (especially Elmslev). Only such a theory is believed to be worthy of the model name. This is very clearly indicated and sufficiently formalized (ideally each model should be done on a computer). [17]

Nowadays, it is even possible to create a 100 percent result if the substitution of synonyms alone can lead to a low level of similarity between the text itself. In natural language, a single sentence or opinion can be expressed in many ways. At this point, of course, the easiest way is to replace the words in the sentence.

1. Alisher Navoi's contribution to the development of the Uzbek language and its formation as a literary language is invaluable. [11]

2. OF COURSE, (THE TRUTH, DOUBT) The contribution of the great thinker-poet, scientist, encyclopedist Alisher Navoi in the development of the Uzbek language and its formation as a literary language is invaluable.

3. Alisher Navoi's contribution to the development of the Uzbek language and its formation as a literary language is invaluable.

4. Depending on the number of cards collected at the end of the lesson, the winning team is determined. Active students are evaluated.

5. Depending on the number of cards collected at the end of the lesson (ALSO BEHAVIOR OF THE LESSON), the winning group is determined. Active students are evaluated.

6. Depending on the number of cards collected at the end of the lesson, (ACTIVE STUDENTS ARE EVALUATED) the winning group is determined.

7. So, teacher, don't be a book lover in the first place! BUT reading is a very general concept.

2. At the stage, the exact same words are compared.

1. 1) Lesson 2) At the end 3) Collected 4) Cards 5) Depending on the number 6) 7) The winner 8) The group 9) is determined. 1) Active 2) Students 3) Assessed.

2. 1) Lesson 2) At the end 3) Collected 4) Cards 5) Depending on the number 6) 7) The winner 8) The group 9) is determined. 1) NOFaol 2) students 3) assessed.

In this case, the first sentence is exactly the same, 2 of the 3 words in the second sentence are similar, but one is completely different, so the sentence is completely different. That is, although the sentence is a logical error, the sentence it expresses is an active antonym in the first and an inactive antonym in the second.

3. Synonymous words are checked at the stage. Based on the existing list, linguistic synonyms are abbreviated, coded S1., S2 and so on, and included in the appendix of our research work.

S1.Abadiy [a], mangu [mo'g'], umrbod [a+f-t], toabad [f-t+a], ilalabad

S2.Abadiylashtirmoq , mangulashtirmoq

S3.Abas [a], foydasiz , behuda [f-t]

S.4Abdol [a], xudojo'y [f-t], dindor [a+f-t]

S5.Abgor [f-t afgor], nochor [f-t], xarob [a]

S6.Abjir , epchil, chaqqon

S7.Abjaq bo'lmoq, majaq bo'lmoq, dabdala bo'lmoq

S8.Ablah [a], yaramas, razil [a]

S9.Ablahona, ablahlarcha

S10.Ablaq [a], chipor, chavkar. In polysemous words, semantics is unique. It can be in a different synonymous sequence with each meaning. For example, the word ending is combined into three synonymous lines with three meanings:

a) to finish, to do, to accomplish. The word is synonymous with the concept of "completion".

b) finish, complete, serve, freeze. In this line, the word is synonymous with the concept of "work task, task performance, accomplishment."

d) finish, sop, lose. The word is synonymous with the concept of "complete annihilation." [9] There are many such words in our language, and it is necessary to mark them separately and form a list with the words that may come.

4. At the stage, the synonymy of phrases is checked in the form of phrases, words and phrases, which are also specially coded (IIBS phrase, SIBS) and a list is attached.

SIBS1. Unanimously - unanimously

SIBS WITH THE WORD = a synonymous phrase

SIBS: GROWING, DEVELOPING, GROWING.

In the current education system, the use of information and communication technologies in accordance with the requirements of the State Education Standard, the formation of critical analysis, project and teaching development, problem-based learning, modules, cases, group work is developing, gaining momentum.

Any model, including a linguistic model, must have a formality specific to literary language. A model is recognized and considered formal if it can clearly show the initial objects that link its statements and the rules for processing them (rules for forming or separating new objects and statements). Ideally, since any formal model is a mathematical system, in a sense, the concept of formality is equivalent to the concept of mathematics, precision, or uncertainty. Formality, clarity, uncertainty are the characteristics of the language in which this theory is presented. By itself, this feature does not ensure that formal theoretical predictions are consistent with objective experimental data. The accuracy of a theory depends on the validity of specific experiments that are able to confirm or refute it. The formal model is linked to the experimental data through this or that interpretation. Interpreting a model means showing the probability or strict rules for substituting objects in a particular subject area, such as language, instead of model objects (symbols). [15] That is, in the formation of language, it is important that the rules of language are manifested in constant uniformity and mathematical precision.

REFERENCES

1. Абжалова М. Таҳрир ва таҳлил дастурларининг модуллари. Т.2020.
2. Safarov.Sh. Kognitiv tilshunoslik. Samarqand. 2006-yil
3. Маслова В. Когнитив тилшунослик. (ўқув қўлланма) Самарақанд – 2005
4. Сайфуллаева Р. Синонимия повествовательных предложений в современном узбекском языке. кан.фил.наук. автореферат, Т,:1982
5. Ҳожиёв.А. Тилшунослик терминларининг изоҳли луғати. Тошкент.ЎЗМЭ. 2002 йил, 88-бет
6. Ҳамроев. М. А. Она тили (ОТМ учун дарслик) Тошкент 2007.64
7. Abdullayeva M va boshq.Qisqacha falsafa lug'ati. Toshkent. 2004-yil
8. Nazarov Q. Falsafa .Toshkent. 2000-yil
9. Sayfullaeva R., Mengliyev B.va boshq. Hozirgi o'zbek adabiy tili – Toshkent. 2010-yil.
10. Бозоров.О Ўзбек тилида даражаланиш. Ф.ф.д.дисс..... Тошкент 1997.
11. Til va adabiyot ta'limi. 2017.10-сон.
12. Lex.uz
13. Ziyouz.com
14. Antiplagiat.ru
15. <https://habr.com/ru/post/341148/>
16. <https://habr.com/ru/post/341148/>
17. <https://xn----7sbbaar5acc1ard1a0beh.xn--p1ai/kak-obmanut-antiplagiat-v-2021>