PPSUTLSC-2024
PRACTICAL PROBLEMS AND SOLUTIONS TO THE
USE OF THEORETICAL LAWS IN THE SCIENCES OF
THE 21ST CENTURY
TASHKENT, 6-8 MAY 2024

www.in-academy.uz

LINGUISTIC AND SOFTWARE OF LANGUAGE CORPUS

[1]Kholmurodov Davron, [2]Munira Xakimova
[1]student of University of Tashkent for Applied Sciences, Gavhar Str. 1, Tashkent 100149,Uzbekistan
[2]scientific supervisor, senior teacher, University of Tashkent for Applied Sciences, Gavhar Str. 1, Tashkent 100149,
Uzbekistan
dxolmurodov292@gmail.com, munirakhakimova2575@gmail.com

Abstract: In this article, it is discussed about the linguistics of the corpus of languages and the software of each language. Linguistics is the scientific study of language, encompassing the analysis of its structure, use, and meaning. One important aspect of linguistics is the creation and analysis of language corpora, which are large collections of written or spoken texts that are used to study language patterns and usage. Software tools are often used in linguistics for the creation, manipulation, and analysis of language corpora. These tools help researchers search for specific linguistic features, extract data, and analyze patterns in the corpus. Some common software tools used in linguistics include text processing tools (e.g. concordancers, text editors), statistical analysis tools (e.g. software for performing quantitative analysis on linguistic data), and annotation tools (e.g. tools for marking up linguistic features in the corpus). The combination of linguistic theory and software tools allows researchers to gain valuable insights into the structure and use of language through the analysis of language corpora. This interdisciplinary approach helps to advance our understanding of language and how it is used in different contexts

Keywords: Corpuslinguistic, lingvaculturology, ortology, rigor, paradigm, tagmemics, modern linguistics, asymmetry, lexicon, junctive expressions.

1   INTRODUCTION

The Department of lexicology was created in 1934, the main purpose of which is Uzbek it consisted of collecting and collecting words in the language. True, even before 1934 the words are collected. Like Fitrat, Munavwar Qori, Chulpan, Botu, Elbek enlightenment textbook of the Uzbek language in the 20s and 30s, creation of Uzbek language rules in addition, they paid great attention to word gathering. Including Elbæk 1923 On September 5, Turkestan opened a chapter" Dictionary and terms "in which Uzbek raised issues of collecting his words. He wrote of it: "We wanted to open a chapter of dictionaries and terms in a newspaper page.[1]

It is known that the application of computer technology in Corpus Linguistics determination of the period and frequency of use of a word based on the resulting language Corpus, development of the field of terminology, study and analysis of sentence construction, creating an N-language base for translation programs, studying methodology, electronic development of linguodidactics, especially of the possibilities of creating various dictionaries the carteque was switched to the automatic process, and bemisl brought comfort to the surface. When it comes to compiling dictionaries, the usefulness of language corpora, especially with frequency the creation of dictionaries is important in the formation of terminology of a particular field counted. Also a translation base of bilingual or multilingual parallel Corps translator dictionaries using [Boliang Zhang 2020], automatic translation it acts as a major resource in the formation of linguistic bases of its programs. In general, the multipurpose and multifunctional language corpus is the national language of a particular language the body is. At such a Corps base 1) of extremely large volumes of texts belonging to all styles of the language be; 2) linguistic, morphological and syntactic touch of texts; 3) from the necessary requirements for the availability of meta-data of the body is. These requirements are necessary linguistic and linguodidactic of the National Corps ensures that it becomes a technological tool. Educational Corps of the National Corps considered a private type, mainly teaching the language in the educational process [Nina Vyatkina 2020], performing various practical exercises in the educational process, text in the research process use of elements, terminology of fields of science related to education the main tool in such matters as formation, formation of an educational thesaurus is.[2]

2   MATERIALS AND METHODS

a) What is the linguistics?

Linguistics or linguistics is a science that studies languages. There are practical and theoretical types of linguistics, and theoretical linguistics studies the structure (grammar) of a language and its meaning (semantics). It covers grammar — morphology (structure and variation of words), syntax (rules of

PPSUTLSC-2024
PRACTICAL PROBLEMS AND SOLUTIONS TO THE USE OF THEORETICAL LAWS IN THE SCIENCES OF THE 21ST CENTURY
TASHKENT, 6-8 MAY 2024

www.in-academy.uz

attachment of words to phrases and sentences) and Phonology (study of language using abstract sounds). Applied linguistics is primarily concerned with the application in practice of theoretical knowledge learned in linguistics. Applied Linguistics includes the study and teaching of foreign languages. Linguistics," linguistics " is the science of language, its social nature, function, internal structure, classification, laws of practice (activity) of certain languages, historical progress. According to its purpose and function, there are several areas (areas) of linguistics: general linguistics is a field that studies language as a human-specific phenomenon in general, the main task of which is the identification and illumination of the most common sign— features characteristic of world languages; private linguistics is a field that studies some linguistic sign-features; applied linguistics is a direction that develops methods fields such as ethnolinguistics, Psycholinguistics, sociolinguistics are related to the activities of the speaker (person) in society.[3]

Lexicology is the part of linguistics dealing with the vocabulary of a language and the properties of words as the main units of language. The term lexicology is composed of two Greek morphemes: lexis meaning – "word, phrase" hence lexicons "having to do with words" and logos which denotes "a department of knowledge". This, the literal meaning of the term lexicology is "the science of the word". The term vocabulary is used to denote the system formed by the sum total of all the words that the language possesses. The term word denotes the basic unit of a given Language resulting from the association of particular meaning with a particular group of sounds capable of a particular grammatical employment. [4]

b) Corpus Linguistics and English Language Teaching Materials.

Indeed, the computerized version of corpus linguistics started to inspire many researchers when some inconsistencies were found between lexical items and grammatical structures in the corpora and those found in traditional language textbooks that were often based on intuition of the material writers. Furthermore, some of the properties of individual words and phrases, such as their frequency and collocations, were not generally open to intuition, which means that corpus analysis can help the syllabus design process. As Ellis argues, "speaking natively is speaking idiomatically, using frequent and familiar collocations, and the job of the language learner is to learn these familiar word sequences" (1997: 129).

Traditionally, however, this naturally occurring discourse (empiricism approach) did not have any place in language theories because descriptions of language and resulting theories were solely based on intuition (rationalism approach). The suspicion about even the simplest statistical evidence in relation to language dates back to one of the most famous opponents of corpus linguistics, Noam Chomsky, who believed in introspection for any insights into the structural nature of language and argued that no collection of naturally occurring discourse can ever be substantial enough to be a true representation of a language because:

Any natural corpus will be skewed. Some sentences won't occur because they are obvious, others because they are false, still others because they are impolite. The corpus, if natural, will be so widely skewed that the description would be no more than a mere list (cited in Tognini-Bonelli, 2001:51).

In fact, what can be inferred from here can be summarized by Derwing's (1979) theory, which states that "performance" – which is our use of language – does not accurately reflect our internal knowledge of language, which is "competence". From Chomsky's viewpoint, grammar is assumed to be part of competence; hence, corpus-derived data that reflects performance cannot be relevant to questions of grammaticality.

CONCLUSIONS

In a nutshell, what we observed in this overview was a simplified functional spectrum which needs to be expanded by means of a more complex and more complete picture. Although a simple pedagogical description of any grammatical feature is perhaps aimed by any textbook author, there is a certain danger of oversimplifying these grammatical structures at more advanced levels. As stated by Barlow (cited in Romer, 2005: 285), "less frequent patterns are important in moving the language learner from intermediate to more advanced levels of proficiency".

Generally, this type of finding points to the fact that a lot of mismatch between traditional descriptions and actual language usage stems from the fact that the strict interconnection between an item and its environment is more or less ignored. As Kennedy noted (1991), the traditional emphasis on the grammatical paradigm has to be revisited in favor of a more syntagmatic approach to be used contextually

References

[1]   To'ra Mirzayev. "Leksikologiya va leksikografiya bo'limi" Tashkent 2020 3p
[2]   Qarorova Dilsora "Korpus tilshunosligining paydo bo'lishi va rivojlanish tarixi."."Iqro" jurnali.Toshkent, 2020 87-88 p

[3]   M.SH.Shirinova "Ortologiya tadqiqi va uning tilshunoslikda tutgan o'rni".Humanitarian Studies.Buxoro 2020 107-108 p

[4]   Khakimova.M.Kh.   Semantic   feature   of phreasological units of the English language denoting "TIME".

[5]   Khakimova Muira "Neologism as a result of the nomination process" 2021/4 jurnal Nauka i obrazovaniya segodnya№4 (63)2021,стр.51-52,Москва 2021

[6]   Munira   Khabibullayevna   Khakimova   "The correlation   between   language   and   culture"   Oriental renaissance: Innovative, educational, natural and social sciences.Vol.2Special Issue № 20 стр.729-732,Izdatel OOO" Oriental Renaissance"