The American Journal of Social Science and Education Innovations
(ISSN – 2689-100x)
VOLUME 04 ISSUE 10 Pages: 01-05
SJIF IMPACT FACTOR (2020: 5. 525) (2021: 5. 857) (2022: 6. 397)
OCLC – 1121105668 METADATA IF – 8.106
Publisher: The USA Journals

Journal Website: https://theamericanjournals.com/index.php/tajssei

Copyright: Original content from this work may be used under the terms of the creative commons attributes 4.0 licence.

Research Article

THE DEVELOPMENT TENDENCIES OF COMPUTATIONAL LINGUISTICS IN UZBEKISTAN: NLP, MACHINE TRANSLATION, CORPUS LINGUISTICS AND AUTOMATIC TEXT EDITING

Submission Date: October 01, 2022, **Accepted Date:** October 05, 2022,
Published Date: October 12, 2022 |
Crossref doi: https://doi.org/10.37547/tajssei/Volume04Issue10-01

Juraev B. Shokhrukh
State World Languages University, Uzbekistan

Shaymardanov H. Abror
State World Languages University, Uzbekistan

ABSTRACT

This paper presents the recent development tendencies in the field of computational linguistics in Uzbekistan. It aims to address the researchers and research papers in computational linguistic areas such as: NLP, Machine translation, Corpus Linguistics and Text Editing. The article also highlights the new branches of Computational linguistics that has gained much importance in recent years in the country.

KEYWORDS

Computational Linguistics, NLP, Machine Translation, Corpus Linguistics, Text Editing, Artificial Intelligence.

INTRODUCTION

Computational linguistics is one of the new directions in the field of linguistics and is located at the crossroads of the arts and sciences. It also deals with the artificial intelligence that is being integrated in linguistics. We can see these elements in natural language processing (NLP) which is the the main sub field of the Computational linguistics. The researchers state that the subjects and disciplines that are related to the artificial intelligence will prosper and play a vital role in our daily lives as well. Even though Computational

The American Journal of Social Science and Education Innovations
(ISSN – 2689-100x)
VOLUME 04 ISSUE 10 Pages: 01-05
SJIF IMPACT FACTOR (2020: 5. 525) (2021: 5. 857) (2022: 6. 397)
OCLC – 1121105668 METADATA IF – 8.106
Publisher: The USA Journals

linguistics has been constantly studied in the developed countries, this field of linguistics gained more importance in only last two decades in Uzbekistan. Below we look at some of the Uzbek researchers and their scientific works in Computational linguistics such as: H.Arzikulov, A. Mamatov, S. Muhamedova, A. Rahimov, Z. Xolmanova, N. Abdurakhmonova, Z. Haitkulov, N. Ataboev and others.

OBJECTIVES

The paper aims to discuss what researches have been done by local scientists in Computational linguistics and what should be done in the field catch up with the modern trends in linguistics. Another goal of the article is to explore the new emerging branches of Computational linguistics and to draw students' attention to them.

GEOGRAPHICAL AREA

In recent years, the importance of teaching world languages in Uzbekistan is increasing. Its demands are increasing in this regard. It is noted that according to the requirements established by the decisions of the Government of Uzbekistan, young people who know English well will have additional benefits: it is planned to increase the salary of employees who have a certificate of knowledge of foreign languages. This is the reason why people in our country are trying to master world languages perfectly today. In accordance with the President's decision in 2021, the interest in language learning in our country will be determined by the creation of video films, entertainment shows, films and other content. It is worth noting that in 2021-2022, the head of our state decided to increase and encourage the number of schools specialized in foreign languages, to support all educational institutions by equipping them with advanced technologies. This is because today we cannot imagine our life without technical tools, because they have many advantages in all areas. As linguistic branches are increasing, new fields are emerging. Computational linguistics is one of the new directions in the field of linguistics and is located at the crossroads of the arts and sciences. Even though Computational linguistics has been constantly studied in the developed countries, this field of linguistics gained more importance in only last two decades in Uzbekistan. So this paper outlines some of the work that has been done so far and discusses further steps to be taken in order to develop the field in the country.

RESULTS

Computational linguistics emerged as an area of Applied linguistics in the second part of the 20th century after Noam Chomskiy's lecture in MIT (Massachusetts Institute of Technology). It officially appeared as a new field when IBM experimented an influential demonstration of machine translation, which was performed in January 7, 1954. It was developed by the Georgetown University and IBM together, the experiment involved automatic translation of sixty Russian sentences into English. Since then, constant researches have been done by the scientists in famous educational establishments and universities. Uzbek researchers started publishing on this area mostly at the beginning of the 21th century. Computational linguistics is a component of applied linguistics, and it is now is rapidly developing in various aspects. The following areas of computational linguistics are mentioned in the publication papers of A. Mamatov, S. Muhamedova, A. Rahimov:
1) automatic translation is a short-term translation of a given text from one language to another quick translation system inside;
2) text editing - editing a specific text using a computer, identifying and correcting errors;
3) CALL (Computer Assisted Language Learning) -

The American Journal of Social Science and Education Innovations
(ISSN – 2689-100x)
VOLUME 04 ISSUE 10 Pages: 01-05
SJIF IMPACT FACTOR (2020: 5. 525) (2021: 5. 857) (2022: 6. 397)
OCLC – 1121105668 METADATA IF – 8.106
Publisher: The USA Journals

language teaching automation - computer-aided foreign language reading system providing practical applications and automatic evaluation with knowledge mechanisms; 4) computer lexicography (creating electronic dictionaries) - a specific language database of words in the lexicon and special programming languages programs (computer dictionaries) that have an operating system using the field engaged in creation; 5) creation of an artificial intelligence system; 6) NLP - natural language processing; 7) hypertext technology; 8) online teaching system (in virtual classrooms and laboratories) - computer language didactics; 9) information search system

D.O'rinboyev grouped the areas of computational linguistics as follows: "Computer Science is a new field, but so far, it is short over time, different directions have emerged. Machine language based on natural languages creation, expansion of user demands and needs is this direction increased. Fields of computer science are conditional are grouped as follows:"

1. Directions for molding and programming:

- formulating communication using a computer;
- computer modeling of the plot structure;
- imagine the text at the level of a large text (hypertext). reach;
- information - search system.

2. Directions for language teaching:

- computerization of language teaching processes;
- machine translation.

3. Approaches to text analysis:

- automatic editing;
- statistical studies;
- computer lexicography.

According to the professor's groupings, Z.Xolmanova divided these areas in the list into two groups. These areas of computer linguistics are practical and theoretical can be grouped in terms of importance as follows:

1. Practical directions:

- Machine translation.
- Automatic editing.
- Computerization of the language teaching system.
- Statistical studies.
- Computer lexicography.

2. Theoretical directions:

- Natural language processing.
- Modeling the plot structure.
- Modeling the dialogue with the help of a computer.
- Providing text in hypertext technologies

Corpus Linguistics. Corpus Linguistics is one of the main streams of Computational linguistics. There are some researches that have been conducted in Tashkent State Uzbek Language and Literature University. A group of researchers have started to work on creating parallel corpus of the Uzbek language. They already launched the first version of the corpus publicly and now working on improving and enriching Uzbek corpus at the university. An international conference "Turklang 2018" was held at the university to gather researchers who are

The American Journal of Social Science and Education Innovations
(ISSN – 2689-100x)
VOLUME 04 ISSUE 10 Pages: 01-05
SJIF IMPACT FACTOR (2020: 5. 525) (2021: 5. 857) (2022: 6. 397)
OCLC – 1121105668 METADATA IF – 8.106
Publisher: The USA Journals

conducting work on Corpus Linguistics from Turkey, Kazakstan, Kirgizistan. The main purpose was to get experience from them on creating corpus Turkik languages.

Machine translation. Machine translation is another branch of Computational linguistics. Recently more and more researches have been conducted in this field as well. For example, the PhD dissertation topic of computational linguist N. Abdurakhmonova is "Linguistic support of the program for translating English texts into Uzbek (on the example of simple sentences)" in Tashkent State Uzbek Language and Literature University. Z. Haitkulov's PhD dissertation theme is related to the issues of the Machine translation "Fundamentals of creating contextual programs of automatic translation" in Uzbekistan State World Languages University.

Automatic Editing. One area of study in computational linguistics is automatic text editing softwares. This sphere is as important as other fields of computational linguistics such as machine translation, automatic speech recognition, Natural language processing. Scientists have recently created a diverse type of text editor programmes. There have been researches in the field of automatic editing in Uzbek linguistics. Scientists group under the leadership of Professor H.Arzikulov established "Linguistics Engineering Center" at Samarkand State Institute of Foreign Languages. Researchers conducted projects on the issues of computer processing of Official language texts in this center. This is a word processing computer creating programs, collecting texts for all levels of the Uzbek language, bringing it to a specific system, creating an algorithm, processing them statistically in such fields as providing, creating an English-Uzbek machine translation program scientific-research works have been carried out. Uzbek text processor automatic system "Uzlington", Latin text into Cyrillic and Krill "Spellchecker", which converts from the alphabet to the Latin alphabet, with the help of a computer computer programs for proofreading and correcting spelling mistakes was one of the first to start creating.

CONCLUSION

There is a shortage of the specialists and the research papers in this field even though some research and educational establishments have succeeded in computational linguistics. The only program that exists in computational linguistics is the masters' program in Tashkent State Uzbek Language and Literature University. Doctoral papers have been made in the other spheres of the computational linguistics in Tashkent State Uzbek Language and Literature University. But there are still a lot to do to catch up with the world standards in the field and to cover these limitations. So there will be a need to have researchers in computational linguistics and programs that train these kinds of specialists. There is also a need to open Bachelor programs that prepares specialists in in computational linguistics.

REFERENCES

1. Po'latov A. (2011) "Computational Linguistics", Sharq Publishing, Tashkent, p.p: 92-94
2. Boyarskiy K. (2013) "General issues of Computational Linguistics".
3. Sankt-Pitersburg p.p: 56-58
4. Xolmanova Z. T.(2020) "Computational Linguistics", Asian Book House, Tashkent, p.p: 24-30
5. Rahimov A. (2016), "Computational Linguistics and Language Science". Tashkent.
6. Mavlanova.T, Ataxanova. Sh (2020) "The role of computer linguistics in the field of philology", Young scholar-Moscow.

The American Journal of Social Science and Education Innovations
(ISSN – 2689-100x)
VOLUME 04 ISSUE 10 Pages: 01-05
SJIF IMPACT FACTOR (2020: 5. 525) (2021: 5. 857) (2022: 6. 397)
OCLC – 1121105668 METADATA IF – 8.106
Publisher: The USA Journals

7. Abdurakhmonova. N, (2018) "Linguistic support of the program for translating English texts into Uzbek (on the example of simple sentences)", PhD Dissertation, Tashkent.