JOURNAL OF ADVANCED LINGUISTIC STUDIES
VOL. 11, NO. 1, JAN-JUN 2024 (ISSN 2231-4075)

Analysis of the Development of Bilingual Thesaurus for the English and Uzbek Languages

NARGIZA SHAMIEVA
National University of Uzbekistan, *Uzbekistan*

ABSTRACT

WordNet is as a widely used lexical database that organizes words into sets of synonyms called synsets, linked by semantic relations such as hyponymy, hypernymy, and meronymy. This paper presents the development of Bilungual Thesaurus for the English and Uzbek languages, referred to as UzWordNet, aimed at facilitating natural language processing and semantic analysis tasks. The thesaurus includes a wide range of lexical entries, each annotated with part-of-speech information and linked to related synsets. EngUzWordNet provides a valuable resource for researchers, developers, and language learners interested in the semantic structure and relationships between words in English and Uzbek. Its development involved collaboration between linguists, lexicographers, and language experts to ensure accuracy and relevance to both languages' linguistic characteristics. This article discusses the methodology, challenges, and future directions for EngUzWordNet, highlighting its potential impact on various linguistic applications.

Keywords: UzWordnet, synsets, hyponomy, hypernymy, meronymy, semantic analysis.

INTRODUCTION

Linguists and psychologists at Princeton University conducted research on creating the WordNet lexical database. At that time,

it was also referred to as a "lexical ontology." The project was inspired by experiments in cognitive intelligence, led by psycholinguist George Miller, aimed at understanding human semantic memory. According to their findings, WordNet contained approximately 95,600 word forms (51,500 words and 44,100 word combinations) with 70,100 meanings provided with synonyms. Approximately 57,000 of these were considered verb forms, with meanings assigned to 47,000 of them. The estimated amount of articles is due to the fact that the indicator is constantly increasing and changing in the online system.

WordNet differs from other conventional dictionaries in that it organizes words not only by nouns, verbs, adjectives, and adverbs but also by their synonyms and hyponyms. WordNet, as described by Fellbaum (1998), serves as a lexical database for psycholinguistic research. This system, developed at Princeton University, is used in linguistic engineering research. WordNet is structured based on synonyms and hyponyms for nouns and troponyms for verbs, as well as antonyms for adjectives. Each word's meaning is provided in a synset (synonym set) with hierarchical and semantic relationships indicated.

METHODS

A significant barrier to the development of linguistic ontologies and conceptual modeling in a new language is the absence of comprehensive lexicographic resources containing detailed and relevant linguistic data. The success of the Princeton WordNet has led to the emergence of wordnets and similar projects for various languages, including multilingual wordnets. In the construction of wordnets, developers often employ the Expand Model, which involves translating available wordnets that establish linguistic relations between items and existing synsets in a source language into equivalent synsets in a target language using bilingual dictionaries.

Most of the wordnets in existence today have been created by translating the Princeton English WordNet. However, an alternative approach, though more laborious and time-consuming, involves compiling synsets and mapping semantic

relations between word senses directly from the data of the language for which the wordnet is being developed. This approach requires access to high-quality dictionaries of synonyms and other semantic resources. The feasibility of developing wordnets largely hinges on the availability of bilingual dictionaries or comprehensive descriptions of the semantic structure of a language. In the case of Uzbek, the lack of extensive English-Uzbek dictionaries (existing ones are limited in scope and primarily used for educational purposes) renders the use of the Expand Model in Uzbek wordnet development impractical. Additionally, the absence of Uzbek semantic dictionaries makes it exceedingly challenging to create an original Uzbek wordnet. The objective of this article is to outline a methodology for constructing an Uzbek wordnet based on a lexical resource. This approach involves directly leveraging the data from the thesaurus, particularly the collection of synsets and the relationships between them. By utilizing this method, it becomes feasible to develop an Uzbek wordnet without relying on extensive bilingual or semantic dictionaries.

The development of wordnets for the Turkish language has been an important endeavor, with two notable projects standing out: the Turkish Wordnet within the BalkaNet project and the KeNet project. The first project, developed at Sabancı University as part of the BalkaNet initiative, represents a significant contribution to the field of computational linguistics and lexical semantics. The BalkaNet project aimed to create wordnets for several languages spoken in the Balkan region, including Turkish. The project adopted a combination of expand and merge approaches to build the Turkish Wordnet. The expanded approach involves translating existing wordnets, particularly Princeton WordNet, into the target language, in this case, Turkish. This approach allows for the transfer of synsets and relationships from the source language to the target language. The merge approach involves merging synsets from different languages that share common topics or concepts. This helps in creating synsets that are typical for each of the BalkaNet languages, including Turkish.

The Turkish Wordnet developed as part of the BalkaNet project contains approximately 15,000 synsets. These synsets include many synonyms for common Balkan topics, reflecting the linguistic and cultural diversity of the region. The inclusion of synsets typical for each BalkaNet language ensures that the Turkish Wordnet is tailored to the specific linguistic and semantic characteristics of Turkish.

One of the key strengths of the Turkish Wordnet within the BalkaNet project is its coverage of diverse topics and concepts. The inclusion of synonyms for common Balkan topics enhances the richness and depth of the Turkish Wordnet, making it a valuable resource for researchers and developers working on Turkish natural language processing tasks. The project's focus on creating synsets typical for each BalkaNet language also ensures that the Turkish Wordnet accurately reflects the semantic nuances of Turkish.

Furthermore, the development of the Turkish Wordnet within the BalkaNet project has contributed to the overall advancement of wordnet research and development. By adopting a combination of expand and merge approaches, the project has demonstrated a novel and effective method for building wordnets for languages with limited resources. The success of the BalkaNet project has paved the way for future research and development in the field of computational linguistics and lexical semantics, particularly for languages spoken in diverse and multilingual regions like the Balkans.

With regard to Uzbek language, the Uzbek English Wordnet Thesaurus is a lexical database that offers an organised list of words and their English equivalents. It is a useful tool for people who want to be able to interpret and comprehend Uzbek and English more successfully. The Wordnet Thesaurus organizes words into synsets, which are groups of synonymous words that represent a single concept. Each synset contains a set of words that are semantically related, allowing users to explore the relationships between different terms in both languages.

People can increase their vocabulary, sharpen their language abilities, and help people communicate across language barriers by using the Uzbek English Wordnet Thesaurus. This resource is

particularly beneficial for language learners, translators, researchers, and anyone interested in gaining a deeper understanding of the Uzbek language.

The thesaurus aims to capture the meanings and relationships between words in both languages, helping users understand the nuances of word usage and facilitating language learning and natural language processing tasks. It can be used by linguists, language learners, and researchers interested in Uzbek and English language semantics.

For cross-lingual research, language instruction, and machine translation applications involving Uzbek and English, the Uzbek-English WordNet Thesaurus is useful. It provides a structured way to explore the meanings and relationships between words in these languages, enhancing our understanding of their semantic structures.

RESULTS AND DISCUSSION

Creating UzWordnet, a comprehensive lexical resource for the Uzbek language, involves several key steps and considerations. The development of bilingual thesaurus requires a thorough understanding of Uzbek linguistic characteristics, the identification and categorization of lexical units, the establishment of semantic relations between these units, and the organization of these units into synsets. This process is essential for capturing the nuances and richness of the Uzbek language and providing a valuable resource for researchers, educators, and developers working with Uzbek language processing applications.

One of the primary challenges in creating bilingual thesaurus is the lack of extensive lexicographic resources for the Uzbek language. While there are some dictionaries and language resources available, they may not cover all aspects of the language or provide the level of detail required for a comprehensive wordnet. Therefore, it is essential to carefully select and analyze existing resources to extract relevant lexical information and ensure the accuracy and completeness of bilingual thesaurus.

At the stage of the creation process of bilingual thesaurus, a team of linguists, lexicographers, and language experts compile a list of lexical units in Uzbek. This list should include nouns, verbs, adjectives, adverbs, as well as their various forms and inflections. Each lexical unit should be assigned a unique identifier and categorized based on its grammatical properties and semantic characteristics.

Once the list of lexical units is compiled, the next step is to establish semantic relations between these units. This involves identifying synonyms, antonyms, hypernyms, hyponyms, meronyms, and other semantic relationships that exist between lexical units. For example, in bilingual thesaurus, the word “kitob” (book) and this word also can be used as adjective “kitob” and would be linked to synonyms such as *nashr* (publication), as well as hypernyms like *ma'lumotnoma* (encyclopedia) and hyponyms like “roman” (novel).

To ensure the accuracy and consistency of semantic relations, it is essential to consult existing linguistic resources and conduct thorough research on the semantic properties of Uzbek lexical units. This may involve analyzing corpora, conducting surveys, and consulting language experts to gather information about the usage and meaning of words in different contexts.

Organizing lexical units into synsets is another crucial step in creating UzWordnet. Synsets are groups of words that are synonymous or semantically related and represent a single concept or idea. For example, in UzWordnet, the synset for *yigit bola* (boy, fellow or guy) and here we should note that choosing appropriate word in English will be done according to context and also this word would include synonyms such as *bolajon* (child) and *o'g'il bola* (son), as well as related concepts like *bola* (child) and *o'g'il* (son).

Creating bilingual thesaurus, a lexical resource that includes Uzbek and English adjectives, involves several steps to ensure accuracy and usability. This process requires careful selection and categorization of adjectives, establishing semantic relations between them, and providing examples to illustrate their usage. Let's explore these steps in detail:

1. **Selection of adjectives**: The first step in creating bilingual thesaurus with Uzbek and English adjectives is to compile a comprehensive list of adjectives in both languages. This list should include adjectives that are commonly used and represent a wide range of semantic categories. Adjectives should be selected based on their frequency of use, relevance to everyday language, and representation of various grammatical forms.
2. **Categorization of adjectives**: Once the list of adjectives is compiled, they should be categorized based on their semantic properties. Adjectives can be categorized based on their meaning, such as size (e.g., big, small), color (e.g., red, blue), quality (e.g., good, bad), or other attributes. Categorization helps organize adjectives and establish semantic relationships between them.
3. **Establishing semantic relations**: Semantic relations between adjectives can be established based on their meaning and usage. Common semantic relations between adjectives include synonymy (similar meaning), antonymy (opposite meaning), hyponymy (subordinate meaning), and hypernymy (superordinate meaning). Creating a bilingual thesaurus involves carefully mapping semantic relations between words in both languages. For the adjective "smart," we need to consider various contexts and meanings in both English and Uzbek. Here's a structured approach:

    1. *Understanding "smart" in English*
    In English, "smart" has several meanings, including:
    a. Intelligent
    b. stylish or elegant
    c. quick-witted or clever
    d. technologically advanced

    2. *Establishing Uzbek equivalents*
    For each meaning, we find corresponding Uzbek terms:
    a. Intelligent - *Aqlli, Ziyrak*
    b. Stylish or Elegant - *Zamonaviy, Chiroyli*
    c. Quick-witted or Clever - Topqir, Zukkor

d. Technologically Advanced - *Texnologik rivojlangan, Aqlli* (for smart devices)

3. *Contextual usage and examples*

To create a useful thesaurus entry, it's important to provide context for each meaning. Here are examples and translations:

a. Intelligent
   English: She is very smart and excels in her studies.
   Uzbek: *U juda aqlli va o'qishlarida muvaffaqiyatli*.
   English: His smart ideas often solve complex problems.
   Uzbek: *Uning aqlli g'oyalari ko'pincha murakkab muammolarni hal qiladi*.
b. Stylish or elegant
   English: He wore a smart suit to the event.
   Uzbek: *U tadbirga zamonaviy kostyum kiyib bordi*.
   English: The decor was very smart and sophisticated.
   Uzbek: *Bezatilishi juda chiroyli va zamonaviy edi*.
c. Quick-witted or clever
   English: She gave a smart reply to the tricky question.
   Uzbek: *U qiyin savolga topqir javob berdi*.
   English: His smart thinking saved the day.
   Uzbek: *Uning zukkor fikrlashi kunni saqlab qoldi*.
d. Technologically Advanced
   English: They live in a smart home with automated systems.
   Uzbek: *Ular avtomatlashtirilgan tizimlarga ega aqlli uyda yashaydilar*.
   English: He bought a smart TV that connects to the internet.
   Uzbek: *U internetga ulanadigan aqlli televizor sotib oldi*.

4. **Semantic relations in a thesaurus format**: In a thesaurus, semantic relations can be shown through synonyms, antonyms, and related terms.

Thesaurus entry for “Smart”
English: Smart (adj.)
Synonyms: Intelligent, clever, stylish, technologically advanced
Antonyms: Stupid, dull, unfashionable, outdated
Related terms: Brilliant, witty, fashionable, digital
Uzbek: *Aqlli* (adj.)
Sinonimlar: Ziyrak, Topqir, Chiroyli, Texnologik rivojlangan
Antonimlar: Ahmoq, Zohiriy, Zamondan qolgan, Eskirgan
Bog'liq atamalar: Aqllilik, Zamonaviylik, Zukkolik, Raqamli

By capturing these various meanings and their equivalents, we can create a comprehensive bilingual thesaurus entry for adjectives in English and Uzbek. This approach ensures clarity and accuracy for users seeking to understand and use the word in both languages.

When combined with distinct nouns, adjectives can take on whole new meanings; that is, new adjectives are created that, when translated into another language (in this case, Uzbek and English), may not have meanings that are very similar to the original adjective.

Providing Examples: Examples are essential for illustrating the usage of adjectives in context. Examples help users understand how adjectives are used to describe nouns and provide additional information about their meaning. For example, the Uzbek adjective *Aqqli* (smart) can be illustrated with the English and Uzbek usage.

AQQLI					
Oqilona	fikr	Aqqli fikrlar aqlli odamlardan chiqadi.	reasonable thing (sage, good, sound, beneficial, prudent, excellent, intelligent, insightful, sensible, helpful)	That is the most reasonable thing this show has ever heard.	logical
Miyali, kallali	**odam**	Manga o’xshash odamxo’rdan ko’ra	a wise man (educated.	The vast rivers	eminent

		sanga o’xshash kallaliyozuvchidan ko’proq qo’rqadi marazlar. X.Do’stmuhammad. Yolg’iz.	knowledgable, intelligent)	move silently, and a wise man keeps his mouth shut.	
Zamonaviy	**uy, uskuna, qurilma, telefon**	Aynan aborigenlar bumerang nomli o‘roqsimon aqlli qurol o‘ylab topishgan. ['N. Mahmudov, A. Sobirov, Sh. Sattorov, Sh. Toshmirzayeva, D. Mannopova «"ONA TILI"» 2020']	A smart home (electronic, authomated)	A smart home system then crept silently into our existence.	intelligent
Dono	**so’z**	Domla o’zini aqqli so’zlari bilan mashhurdirlar.	wise words (sage insightful wisdom profound sensible enlightened great advisable intelligent smart knowledgeable philosophers beneficial)	These wise words can be applied to other spheres of life, as well.	sage
Farosatli	**ayol**	Oilani saqlab qolish uchun aqlli ayol bo’lish lozim.	intelligent woman(educated. knowledgable, intelligent)	She is a highly intelligent woman, and we can appreciate her viewpoint.	highly educated
Esli, oqil	**bola**	Abdulla esli, aqqli bola, katta odam bo’lib qaytardi. O’.Umarbekov. Odam bo’lish qiyin.	a bright young man (intelligent educated, knowledgeable, clever, bright, intellectual, brilliant, savvy, gifted, skilled, sharp, talented, competent, shrewd, creative, capable, logical,	He's going to be a bright young man in the future.	outstanding

			highly intelligent, sophisticated, thoughtful, able-bodied, enlightened literate, academically, brainy, calculating wise, highly, educated, perceptive)		

Integration of Uzbek and English adjectives

To create a bilingual thesaurus, Uzbek and English adjectives should be integrated into a single database. This database should include information about each adjective, such as its part of speech, meaning, and semantic relations. Adjectives should be linked to their respective synsets and examples to provide a comprehensive resource for users.

According to our research we fint out lots of synonims to the adjective SHARP in Uzbek as well as in English. Originating from the uzbek language, we find the English translation by distinguishing the adjectives that indicate the characteristics of the sign. We used online and other dictionaries in our work. Especially REVERSO, the uzbekcorpus.uz platform, and Nizomiddin Makhmudov's "O'zbek tilinining izohli lug'ati". We compared any selected adjective with a word (noun) next to it. For example, we find the word O'TKIR (sharp) in explanatory dictionaries, and looking at its combinations with nouns, we can see that it comes up with a new meaning adjectives every time.

o'tkir tig' (*keskir*) - jagged peak (cutting)
o'tkir hid (*kuchli*) - pungent stench (stinky)
o'tkir qalam egasi (*iste'dodli*) - tempting voice (talented)
o'tkir so'z (*ta'sirli*) - a sharp word (impressive)
o'tkir ong (*aqlli*) - keen mind (clever, smart, intelligent, knowledgeable)
o'tkir tansiqlik (*muhtoj*) - a severe lack (deficiency)
o'tkir xis (*kuchli*) - strong sense (radical)
o'tkir ko'z (*ziyrak*) - good eyesight (clairvoyant, keen-sighted)
o'tkir nigox (*teshuvchi*) a keen eye (meaningful)

o'tkir kofe (*achchiq*) bitter coffee (acerbic)
o'tkir ovoz (*qattiq*) - loud voice (strong)

This way we can prove that not only basic synonyms to the adjective could be their alternatives but while combinating with a certain word (noun) they acquire absolutely new meaning and they also shall be synonyms to the basic adjective.

O'TKIR					
keskir	Tig'	Fransuz Polineziyasining geografik jihatdan muborak bo'lagining markazida siz Otemanu tog'ining qirrali cho'qqisini topasiz, uning chekkasida mayda orolchalar va rang-barang dengiz hayoti bilan to'lib-toshgan marjon rifi joylashgan.(Wikipedia)	jagged peak	At the heart of this geographically blessed slice of French Polynesia, you'll find the jagged peak of Mount Otemanu, with tiny islets and a coral reef teeming with colorful marine life sitting at its fringes.	Cutting
Kuchli	hid	Gul chivinlarni o'ziga tortadigan o'tkir hidni chiqaradi.	Pungent stench	The flower gives off a pungent smell that attracts flies.	Stinky
Iste'dodli	Qalam egasi	Darhaqiqat, qora tanlining yagona quroli uning vasvasaga soluvchi ovozi edi.	tempting voice	The Dark Lord's only weapon, in fact, was his tempting voice.	Talented
Ta'sirli	So'z	Jangovar filmning qahramoni nafaqat aniq suratga olish, balki o'tkir so'z hamdir.	a sharp word	The hero of the action movie is not only a precise shooting, but also a sharp word.	Impressive
Aqlli	Ong	Qo'rqmaslik va o'tkir aqlingiz sizga ko'p narsaga erishishga yordam beradi.	keen mind	Your fearlessness and keen mind will help you achieve a lot.	Clever, smart, intelligent, knowledgeable
Muhtoj	Tansiqlik	Oziqlanishning jiddiy etishmasligi qurbonlarning sog'lig'iga qanday ta'sir qilishi	a severe lack	One can easily imagine how a severe lack of nutrition could	Deficiency

		mumkinligini osongina tasavvur qilish mumkin.		affect the health of the victims.	
kuchli	Xis	Ujuda kuchli hazil tuyg'usiga ega va u jozibali.	strong sense	She has a... has a very strong sense of humor, and she's attractive.	Radical
Ziyrak	Ko'z	Ularning juda sezgir burunlari, yaxshi ko'rishlari va ovga bo'lgan tabiiy ishtiyoqi bor.	good eyesight	They have very sensitive noses, good eyesight and natural enthusiasm for the hunt.	Clairvoyant, keen-sighted
Teshuvchi	Nigox	Ushbu rolda siz tafsilotlarga va mukammal muloqot qobiliyatlariga ega bo'lishingiz kerak.	a keen eye	In this role, you should have a keen eye for detail and excellent communication skills.	Meaningful
Achchiq	Kofe	Ertalabki nonushta uchun men achchiq qaxva ichishni hush ko'raman.	Bitter coffee	For the morning breakfast I prefer bitter coffee.	Acerbic
Qattiq	Ovoz	Fuqarolarning bu boradagi baland ovoziga quloq tutilganidan xursandmiz.	loud voice	We are glad that the loud voice of citizens on this issue was heeded.	Strong

Development of user interface

Finally, the bilingual thesaurus database should be integrated into a user-friendly interface that allows users to search for adjectives, explore semantic relations, and access examples. The interface should be accessible online and include features such as search functionality, browsing capabilities, and cross-referencing between Uzbek and English adjectives.

The English-Uzbek Dictionary, published in 2008, is a valuable lexical resource utilized in the creation of bilingual thesaurus. This dictionary is the result of collaboration between the authors and experts at the Uzbek World Languages University and the Uzbek Academy of Sciences. It stands out as one of the largest bilingual dictionaries available electronically, boasting a rich collection of entries.

The 2008 edition of the dictionary contains approximately 40,000 lemmas in the English-Uzbek section and around 30,000 in the Uzbek-English section. Each English word in the dictionary is accompanied by its Uzbek equivalents, presented in a standardized format. For instance, for the English word "sense," the dictionary provides Uzbek equivalents as follows: "sense" [sens] n 1) his, *tuyg'u*, *sezgi*; 2) *aql*, *fahm*, *idrok*, *zehn*.

Here, the numbers represent different senses or meanings of the English word "sense." Additionally, each lemma's entry in the English-Uzbek section of the dictionary includes the major parts of speech associated with the lemma. However, it is important to note that the "connectivity restoration algorithm" used in the creation of UzWordnet focuses on nouns, adjectives, verbs, and adverbs. This algorithm processes the Princeton WordNet (PWN) and its semantic network to generate bilingual thesaurus, utilizing the linguistic information provided in the English-Uzbek Dictionary to enrich the lexical database.

Another example from the English-Uzbek Dictionary could be the word "beautiful": beautiful [byu-ti-fuhl] adj 1) *chiroyli, go'zal*; 2) *ajoyib, hayratli*. In this example, the adjective "beautiful" is translated into Uzbek as *chiroyli* and *go'zal* for the first sense, and as *ajoyib* and *hayratli* for the second sense. The dictionary provides multiple Uzbek equivalents to capture the various nuances and shades of meaning conveyed by the English word "beautiful."

To ensure the usability and accessibility of bilingual thesaurus, it is essential to develop user-friendly interfaces and tools that allow users to easily navigate and search the wordnet. This may involve creating online databases, mobile applications, or desktop software that provide access to bilingual thesaurus and allow users to explore lexical units, synsets, and semantic relations.

In conclusion, the creation of bilingual thesaurus is a complex and challenging process that requires careful planning, research, and collaboration. By following the steps outlined above and leveraging existing linguistic resources, it is possible to create a comprehensive and valuable lexical resource for the Uzbek language that can serve as a foundation for further

research and development in Uzbek language processing applications.

REFERENCES

Abdurakhmonova, N. 2016. The bases of automatic morphological analysis for machine translation. *Izvestiya Kyrgyzskogo gosudarstvennogo tekhnicheskogo universiteta*, 2/38, 12-7.

——. 2012. *Mashina tarjimasining lingvistik asoslari.*

——. 2019. O‘zbek Tilida Wordnet Yaratish Masalalariga Doir. O‘zbekistonda xorijiy tillar” ilmiy-metodik elektron jurnal /journal.fledu.uz/

——. & Abduvaxobov, G. I. 2021. O ‘quv lug ‘atini tuzishning nazariy metodologik asoslari. СЎЗ САНЪАТИ ХАЛИҚАРО ЖУРНАЛИ, 103.

——. & Muhammad, K. 2019. On the tasks of creating a Wordnet in the Uzbek language (uzbek). *O‘zbekiston Xorijiy Tillar* [Foreign Languages in Uzbekistan], 4, 19-27.

——., Alisher, I. & Sayfulleyeva, R. 2022. MorphUz: Morphological analyzer for the Uzbek language. In *7th International Conference on Computer Science and Engineering (UBMK)* (pp. 61-66).

Abduraxmonova, N. Z. 2018. Linguistic Support of the Program for Translating English texts into Uzbek (on the example of simple sentences). Ph.D dissertation.

Agostini, A., Usmanov, T., Khamdamov, U., Abdurakhmonova, N., Mamasaidov, M. *UZWORDNET: A Lexical-Semantic Database for the Uzbek Language*.

Bilgin, O., Çetinoğlu, Ö. & Oflazer, K. 2004. Building a Wordnet for Turkish. *Romanian Journal of Information Science and Technology*, 7/1-2, 163-172.

Çetinoğlu, Ö., Bilgin, O. & Oflazer, K. 2018. Turkish Wordnet. In K. Oflazer & M. Saraçlar (Eds.), *Turkish Natural Language Processing. Springer*. doi:10.1007/978-3-319-90165-7_15.

Fellbaum, Ch., 1998. *Wordnet. An Electronic Lexical Database*. Massachusetts Institute of Technology.

Mengliev, D., Barakhnin, V., Eshkulov, M., Palvanov, B., Abdurakhmonova N. & Khamraeva S. 2023. Dictionary-based medical text analysis in Uzbek: Overcoming the low-resource challenge. In *2023 IEEE Ural-Siberian Conference on Computational Technologies in Cognitive Science, Genomics and Biomedicine (CSGB)* (pp. 85-89).

Nilufar, A. & Gulruhsor, H. 2024. O ‘zbek tili tezaurusini yaratishda leksik birliklarning ifodalanish. In *Konferensiyalar| Conferences*, 1/10, 186-192.

Razieh, E. 2018. KeNet: A Comprehensive Turkish Wordnet and Using it in Text Clustering. PhD Thesis, Işık University.

Sulevmanov D., GatiatullinA., Prokopyev N. & Abdurakhmonova N. 2020. Turkic morpheme web portal as a platform for turkology research. In *2020 International Conference on Information Science and Communications Technologies (ICISCT)* (pp. 1-5).

Ulugbek, T. & Nilufar, A. 2024. Method of developing artificial patterns to analyzing semantic similarity of documents. *Contemporary Technologies Of Computational Linguistics*, 2/22/04, 362-365.

Vladislav, K. & Nilufar, A. 2024. The development of an application for collecting and forming corpora of the Kazakh language. *Contemporary Technologies of Computational Linguistics*, 2/4, 260-265.

Vossen, P. J. T. M. 2002. *EuroWordNet: General Document*.

NARGIZA SHAMIEVA
DEPARTMENT OF FOREIGN LANGUAGE AND LITERATURE,
NATIONAL UNIVERSITY OF UZBEKISTAN, UZBEKISTAN.
E-MAIL: <SHAMIEVANARGIZA5@GMAIL.COM>