

## AUTOMATED RECOGNITION OF PROPER NOUNS IN NER OBJECTS (ANTHROPONYMS)

Madina T. Samatboyeva

Computational linguistics, and Digital technologies,  
Tashkent State University of Uzbek language and Literature  
Tashkent, Uzbekistan  
msamatboyeva@gmail.com

### Abstract

This article discusses proper nouns, their classification, their study within the field of onomastics, and the onomastic layer. Proper nouns differ significantly from common nouns due to their distinctive characteristics. Proper nouns can be categorized into various types, and even within major types, they can be classified into smaller subgroups based on the reasons for their naming. Anthroponyms (personal names) make up the largest category of proper nouns. The article specifically focuses on the identification of proper nouns, particularly anthroponyms, as part of NER (Named Entity Recognition) objects.

**Keywords:** proper nouns, common nouns, onomastics, NER, named entities, anthroponyms.

### Introduction

#### Introduction:

In recent years, the theoretical methods of names and proper nouns within the field of onomastics have been thoroughly studied in linguistics. Several scientific works have been produced, including dissertations and monographs. However, while the linguistic features of names, especially proper nouns, have been explored from a linguistic perspective, their automatic identification in texts within the scope of computational linguistics has not yet been sufficiently researched. When processing Uzbek texts, the identification of proper nouns, such as personal names (anthroponyms) and place names (toponyms), requires attention to their non-dictionary lexical nature and linguistic and grammatical characteristics.

#### Main Section:

Understanding the types and range of proper nouns in a given language helps identify the scope of proper nouns and the onomastic field within that language. This problem is often interpreted in onomastic research using the concept and term “onomastic scope.” For the Uzbek language, the following “onomastic scope” can be presented (Figure 1).

Figure 1. The “Onomastic Scope” of the Uzbek Language

The largest category, encompassing all named objects in the Uzbek language.

- Large category: Contains broad groups like personal names, place names, etc.
- Small category: Contains smaller subcategories like cosmonyms, etc.
- Other categories: Refers to other types of names not included in the two main categories.

Although some studies on onomastics mention the term “onomastic scope” this concept is often considered broad and abstract, referring to the totality of proper nouns within a language. These proper nouns are usually grouped into large categories: toponyms, anthroponyms, zoonymy, and others (Figure 2).

The smaller onomastic fields within this macro scale represent different types and groups of named objects. Some group names may seem unfamiliar at first glance, but their meanings can often be understood by translating their structure (e.g., toponymy – place names, anthroponymy – personal names). There are other philological terms representing groups of named objects in society: demononym – specific names given to devils or spirits; mythonymy – fictional or legendary proper nouns from folk tales and epics; anemonym – proper names given to natural elements like storms, winds, and cyclones.

Figure 2. The Macro-Scale of Uzbek Onomastics Named Macro Categories:

- Anthroponymy
- Zoonymy
- Theonymy
- Toponymy
- Cosmonymy
- Phytomymy
- Chrononym
- Ideonym
- Anemonym
- Chrematonym
- Polityonym
- Ergonym
- Falernonym
- Dignityonym
- Mythonym
- Poreyonym
- Stratonym
- Temeronym

Undefined Macro Categories:

- Festivals, anniversaries, and other ceremonial names
- Major historical events:
- Companies, wars, military campaigns, and other significant historical events
- Political, ideological, and other societal groupings or movements

According to Russian scholar N.V.

Podolskaya, anthroponyms can be divided into several types: names, surnames, patronymics, cryptonyms (nicknames), pseudonyms, andonyms (author's name), patronyms (names reflecting one's lineage), and genonyms (clan or family names).

Before delving into the linguistic meaning of the term "anthroponym," we must first look at the terms onomastics and anthroponymy. Onomastics is the branch of linguistics that studies names. It examines all proper nouns and their lexical meanings.

The onomastic system of the Uzbek language is vast and rich in meaning, representing an interconnected linguistic phenomenon. As the lexical richness of a language consists of specific lexical units (lexemes), the onomastic corpus of a language is made up of proper nouns. These specific names are referred to as "onomastic units," in a similar way to terms like "lexical units" or "speech units" in linguistics. Azerbaijani linguist A. Gurbanov states that "onomastic units" represent the concept of proper nouns within onomastics.

Anthroponymy is the subfield of onomastics that studies personal names. The object of study in anthroponymy is the anthroponym (personal name). Anthroponymy analyzes the function of anthroponyms in speech, changes in an individual's social or familial status, the transformation of names, and the structure and development of anthroponymic systems.

The term anthroponym comes from the Greek words "anthros" (man) and "onoma" (name), meaning "person's name". Anthroponyms refer to special individual names given to people for the purpose of distinguishing them. Human naming has existed as a tradition since ancient times, arising from the need to distinguish and call individuals by unique names due to everyday requirements, cultural practices, religious beliefs, and environmental influences.

#### The Study of Anthroponyms

The study of anthroponyms in Uzbek linguistics has primarily focused on the lexicosemantic features of these names and their cultural, nominative, and linguistic aspects. Scholars like M. E. Rut, A. V. Superanskaya, and N. I. Tolstoy have made significant contributions to this field. Uzbek linguists like S. Ibrohimov, E. Begmatov, and D. Abdurahmanov have also explored these themes in their works, publishing numerous articles and monographs.

In 1965, E. Begmatov defended a dissertation on "Uzbek Anthroponymy," marking the first thorough linguistic and extralinguistic analysis of proper nouns. His research focused on the lexicosemantic features of Uzbek names, the formation of onomastic names, and the history of Uzbek surnames and patronymics.

#### NER Objects and the Study of Anthroponyms Worldwide:

In global computational linguistics and the NLP field, *NER (Named Entity Recognition)* objects have been studied across various domains, such as medicine and search-based systems. For example, in Chinese medical texts, drug names are identified as NER objects using specific models. Researchers have developed a multi-level semantic fusion network for identifying named entities in medical texts. Similarly, in another study, researchers used semi-supervised learning to recognize personal names in comparison between two texts.

---

### Conclusion:

The grammatical field in linguistics is vast, encompassing both morphology and syntax. In terms of morphology, proper nouns are unique lexical categories. However, despite being well-researched, models and programs that automatically recognize these nouns in texts have not yet been developed, emphasizing the need for continued research in this area. For Uzbek, proper nouns, especially anthroponyms, can be accepted as NER objects, given that NER identifies names in a given text, with proper nouns being lexemes not included in standard dictionaries.

### References:

1. Бегматов Э. Ўзбек тили антропонимикаси. Ўзбекистон Республикаси Фанлар академияси “Фан” нашриёти Тошкент. 2013. 35р
2. Бегматов Э. Ўзбек тили антропонимикаси: Филол. фанлари номзоди...дисс. – Т., 1965. 360р.
3. Теория и методика ономастических исследований. р. 9-10.
4. Подольская Н. В. Словарь русской ономастической терминологии. – М. : Наука, 1978.31
5. Бегматов Э. Нафасов Т. Караев С. Ономастика Узбекистана. – Т.: Ўқитувчи, 1989. – 208 р
6. Гурбанов А.М. Азербайжон ономастикаси meseleleri. – Баку,1986.
- 7.Navbahor Nazarova. Antroponimlarning o’rganilishi. Integration of Pragmalinguistics, Functional Translation Studies and Language Teaching Processes. International Scientific Conference, December 18-19, 2021. 110р
- 8.<https://fayllar.org/3-maruza-ozbek-antroponimikasi-haqida-tushuncha-va-uning-asosi.html>
9. Рут М. Антропонимы: размышления о семантике // Изд. Урал. гос. ун-та. – 2001 – No20 [Сер.] Гуманитарные науки. Вып. 4. – С.59-64; Суперанская А. Общая теория имени собственного / Суперанская. А – М.: Наука, 1973. – 365 с.; Толстой Н. Имя в контексте народной культуры /Толстой Н // Язык о языке: Сб. статей. – М.: Языки русской культуры, 2000. – С.101-111; Алиев Г. Лингвистические и экстралингвистические факторы в антропонимии (на материале тюркских языков: азербайджанского, узбекского, казахского, киргизского, татарского и башкирского): Автореф. дис. ... докт. филол. наук / Алиев Г. – Алма-Ата, 1989. – 48 с.; Баскаков Н. Элемент GUL «роза, цветок» в составе каракалпакских имен / Баскаков Н // Ономастика и норма. – М.: Наука, 1976. – С.138-143; Благова Г. Антропонимическая система как проекция космологических и социальных представлений пратюрков /Благова М// Сравнительно-историческая грамматика тюркских языков. Пратюркский язык- основа. Картина мира пратюркского этноса по данным языка. – М.: Наука, 2006. – С.660-754; Гафуров А. Имя и история /Гафуров А. – М.: Наука, 1987. – 220 с.; Жанузаков Т. Лично-собственные имена в казахском языке: Автореф. дис... канд. филол. наук / Жанузаков Т. – Алма-Ата, 1969. – 28 с.; Махпиров В. Имена

- далеких предков / Махпиров В. – Алматы: Ин-т востоковедения МН-АН РК, 1997. – 299 с.; Никонов В. Женские имена в России XVIII века / Никонов В // Этнография имен. – М.: Наука, 1971. – р.120-141;
10. Матвеев А. Ономастика ва ономатология: терминологик этюд // Вопросы ономастики. – Екатеринбург, 2005, №2. – С. 6-7.
11. Qilichev B. Onomastika. O'quv qo'llanma. Vuxoro. 2023. 98
12. Абдуллаев Ф. Киши отларининг қисқариш усуллари //Ўзбек тили ва адабиёти. - 1980. 3-сон. –Б.3- 7
13. Шамсиева М. Кишиларга атаб қўйилган номлар// Адабиётшунослик ва тилшунослик масалалари. -Т., 1962. 4-китоб. –Б. 456-463.
14. Бегматов Э. Ўзбек тили антропонимикаси: Филол. фанлари номзоди...дисс. – Т., 1965. –480 р
15. Бегматов Э. Киши номларининг маъноларига доир// Ўзбек тили ва адабиёти, - 1992. 2-сон. –Б.5; Ономастик сўз яшаш масалаларига доир//Ўзбек тили ва адабиёти. 1965. 3-сон; Номлар ва одамлар. –Т.: Фан, 1966. – 65 б.; Ўзбек исмлари имлоси. –Т.: Фан, 1972. – 308 р.
16. Jintong Sh., Mengxuan S., Zhengya S., Mingda L., Yifan G., Wensheng Z. Multi- level semantic fusion network for Chinese medical named entity recognition. Journal of Biomedical Informatics. 2022. 104p. <https://www.sciencedirect.com/journal/journal-of-biomedical-informatics>
17. Jubo Ch., Chuhan W., Tao Q., Zhigang Y., Yuesong Z., Shuai Y., Jin G., Donghong S., Yongfeng H. Semi-supervised named entity recognition in multi-level contexts. Neurocomputing. 2023. 194p. <https://www.sciencedirect.com/journal/neurocomputing>.