



Representativeness of Mythonyms and Theonyms in Language Corpora

Tlegenova Gulden Bakytkazyevna^{1*}, Rysbergen Kyzdarkhan Kurmashkyzy², Zhanabekova Aiman Abdildaevna²

¹Al-Farabi Kazakh National University, Kazakhstan, ²The A. Baitursynov Institute of Linguistics, Kazakhstan

Abstract The article provides a comparative analysis of the use of mythonyms and theonyms in onomastic subcorps of the world's major languages – English, German, Czech, Kazakh, and Tatar. The research is focused on defining their mention frequency in several textual styles and cultural code analysis which marked these naming units. Using the computer-assisted methods of text searching and passportization, the authors researched the scope and content of the database focusing on motivational analysis of the terms mentioned. Particular attention was paid to the comparison of pragmatics and cultural semantics of theonyms including their use in phraseologisms. The research results emphasized the sematic and structural convergences and differences between the languages by having beneficial effect on better understanding of expressions and thinking ways in various linguistic and cultural contexts.

Keywords: Applied linguistics, Mythonyms, Theonyms, Language corpus, Onomastics

1. Introduction

In the contemporary world, corpus linguistics plays an integral role in exploring language phenomena. It allows for analyzing and interpreting linguistic symbols and texts, enriching the database and improving the accuracy of linguistic studies. The opportunities of corpus linguistics were significantly expanded by the development of technology and text editing systems and specifically, by the consumption of high-powered searching systems and methods of statistical analysis. However, in studies on corpus linguistics, onomastic vocabulary still stays relatively underexplored in spite of the wide range of research phenomena. It is caused by its specificity and variety of forms, which makes its analysis remarkably interesting and crucial. The onomastic lexicon is the most favored data for corpora studies. One of the significant works in this direction is the study of Motschenbacher (2020). The author, noting that corpus linguistics is still an understudied methodology in onomastics, proposes a theoretical understanding of onomastics from the point of view of corpus linguistics, establishing a closer

<https://doi.org/10.22034/ijscsl.2024.2017262.3288>

***Corresponding Author:**

Tlegenova Gulden
Bakytkazyevna
gulden_20.88@list.ru

Received: November 2023

Revised: January 2024

Accepted: January 2024

Published: February 2024

© 2024 Bakytkazyevna,
Kurmashkyzy, and
Zhanabekova.

This is an open-access article
distributed under the terms of
the Creative Commons
Attribution License (CC BY).

connection between them. When analyzing the peculiarities of cultural connotation, or, more precisely, cultural-national connotation, it is postulated that the system of images fixed in the phraseological composition of a language serves as a kind of "niche" for the cumulation of worldview and is somehow related to the material, social, or spiritual culture of a given linguistic community, and therefore can testify to its cultural-national experience and traditions (Telia, 1996). Thus, Simpson and Swales call corpus linguistics a technique or technology for creating and analyzing a corpus (McEnery et al., 2006). This intricate interplay between language and culture is also evident in theonymy. Theonymy, as noted by a number of researchers, is characterized by the peculiarity of synonymy – in the field of theonyms, it is synonymy, which does not rely on the connection with the concept but is based on the identity of objects. Superanskaya and Suslov (2008) describe this phenomenon as polyonymy or plurality. An example of polyonymy is the extensive designations for God, such as Lord, Creator, Almighty, etc.

The study of the onomastic lexicon, in particular mythonyms and theonyms, is an important aspect of linguistic research. These elements of the language have a special semiotic and constructive function in the text, often acting as a semantic dominant. Their analysis allows for a deeper understanding of cultural and national connotations fixed in the phraseology and ways of thinking of different linguistic communities. Mythonyms and theonyms become key to understanding cultural heritage, the emotional coloring of texts, and the metaphorical level of language.

The purpose of this research is to conduct a complete examination of mythonimic and theonymic markings in several language corpora (English, German, Czech, Kazakh and Tatar). Several criteria influence the selection of certain language corpora for investigating the representativeness of mythonyms and theonyms. These may include linguistic diversity, cultural contextualization, and comparability and comparison. Thus, the selection of these language corpora is motivated by a goal to increase the study's complexity and discover various features in the use of mythonyms and theonyms across languages and civilizations. The study seeks to identify and explain the meanings of these aspects, as well as to comprehend their significance in language use and cultural symbols. Special emphasis is placed on the comparative analysis of text pragmatics, statistical definitions, and cultural codes, which will help us to gain the better understanding of the impact of mythonyms and theonyms on the diversity of linguistic expressions and cultural settings.

2. Theoretical Framework

The study of cultural-linguistic codes in separate layers of widely used vocabulary, such as phraseology and metonymic vocabulary, is a relatively new task in world linguistics. At present, this issue is investigated in a fragmentary way. This problem is somewhat new to the field of world linguistics, and the cultural-linguistic codes of separate strata of the commonly used lexicon (phraseology, metonymic lexicon, etc.) are currently studied fragmentarily. The scientific and theoretical basis of onomastics was supplemented by the studies of scholars with a linguocognitive orientation, which became weighty theoretical, scientific, and methodological initiatives in the following ten years. On the basis of these works, as well as the study of national corpora of Turkic, English, and Russian languages, a number of scientific works were created (e.g., Chomsky, 2006; Gak, 1999; Radlov, 1899-1911; Superanskaya & Suslov, 2008; Telia, 1996).

In the world of onomastics, there exist problems with studying religious and mythological codes of different languages, as well as systematization, and modernization of their semantics in the context of modern linguoculturology. Scientific and linguistic interpretation of cultural and historical codes of onomastic names enable us to comprehend the traditional world and know the people "from the inside" through spiritual and ethical values, and folk ecological cognition. In turn, this should arouse the interest of the younger generation and people learning the language. Information about any people, nation, or ethnos, including spiritual and moral values, national spiritual priorities, and stereotypes expressed in names, can open new sources of knowledge about its traditions and customs.

The study of corpus onomastics employs powerful methods in the examination of proper names. One main technique is the simultaneous use of the frequency method and the method of qualitative analysis, which allows for comprehensive revelation of the rich nature of linguistic phenomena and

obtaining valuable scientific results. In one of the chapters of his paper, Motschenbacher (2020) appropriately uses the phrase "corpus linguistic onomastics", and proposes a strong integration of four types of corpus linguistic analysis related to onomastic names: frequency analysis, concordance analysis, word combination analysis, and keyword analysis.

Simultaneously, increasing the corpus of theonyms is carried out by a number of methods, including the principle of metaphorical transfer. The nomination is based on the assumed properties of God: Sinless, Immortal (based on the denial of human qualities) or Omnipotent, Omniscient (indicating the absolute, special completeness of the attribute). Christian theonyms can act as independent mythonyms (Jesus Christ, Sabaoth, Mary) or as mythoappellatives (King, Sovereign, Protector, Advocate). Theonomic units can be mythosymbols or mythologemes, indicating their cultural significance. As such, they are characterized by their globality and universality. In a broader linguistic context, this methodological integration aligns with the broader trends in linguistic research where different scholars employ diverse methodologies for a deeper understanding of linguistic and cultural dynamics. Through a comparative analysis of corpora of different periods in the works of a Finnish scholar language changes and peculiarities of historical development related to personal names were identified (Dang & Vartiainen, 2019). German scholars, on the other hand, incorporate geographical information into their examination of historical corpora in a cultural context (Tse, 2003). Furthermore, foreign specialists recommend combining corpus linguistics and geographic information systems (Baker et al., 2019).

In the onomastic subsystem mythonyms and theonyms are represented by ethnomarker units in the form of a single word, phraseology, proverb, saying, precedent name, precedent text, or concept. The study of the components of mythological and theological images of the world is currently being intensively conducted within the framework of ethnolinguistics, linguocultural studies, linguo-cognitive, and cognitive-discursive directions. The features of archaic and modern mythological and theological consciousness are considered on the basis of folklore and artistic discourses. Lexicographic, dialectal, historical sources, precedent texts, and names formed in their time in the image of the people's world are taken as research materials. These directions of research are noted by such authors as McEnery and Wilson (2001) and McEnery et al. (2006).

When discussing mythological and theological data, the notion of representativeness gains significance, as there is a need for a comprehensive portrayal of linguistic elements within cultural and religious contexts. In general, representativeness is considered by Biber (1993) as a representation of a wide range of functional styles and genres in a corpus of texts. In turn, Baker (2012) writes that the concept of representativeness is closely related to the concept of validity or correspondence of the obtained data to the real state of the language in a given sphere of use (Contemporary Corpus Linguistics, 2012), at the same time, as researchers believe, full representativeness in corpora is unattainable and impossible. As Krasina and Perfilieva (2015) note, the verbal image, and complex figurative structures are governed by deep associative links that are found at different levels of the dynamic structure of an artistic text. These considerations underscore the complex challenges involved in achieving representativeness in linguistic corpora, emphasizing the interplay between language and its multifaceted contexts.

3. Methodology

3.1. Materials

3.1.1. Corpus for the English Language

The British National Corpus is a 100 million-word corpus of texts containing samples of written and spoken British English from a wide range of sources. The corpus covers British English of the late twentieth century, represented by a wide variety of genres, and is intended as a sample of typical spoken and written British English of the time. Three publishers collaborated on the BNC project "Oxford University Press as lead co-author, and Longman and W. & R. Chambers", two universities "Oxford University Press and Lancaster University Press" and the British Library. The creation of the BNC began in 1991 under the leadership of the BNC consortium and was completed by 1994. No new

examples were added after 1994, but the BNC underwent minor revisions before the second "BNC World, 2001" and third "BNC XML Edition, 2007" editions.

3.1.2. *Corpus for the German Language*

The *corpus for the 20th-century German* dictionary was assembled by the Berlin-Brandenburg Academy of Sciences. It includes several sub-corpora: Fiction (26%), newspapers (27%), scientific literature (22%), non-fiction texts (20%), and oral texts (5%). In summary, the corpus includes 106 million words and more than 200,000 articles. Time periods and text types can be selected separately in the corpus. The corpus is also linked to other corpora, e.g. the 21st Century German Corpus.

3.1.3. *Corpus for the Czech Language*

The *Czech National Corpus* (CNC) Project was founded in 1994 to create and make available extensive language data for teaching and research in the form of electronic corpora. It currently provides access to over three billion words organized into synchronic and diachronic, oral and written, parallel and monolingual corpora. The CNC project is mainly managed by staff from two UK FF institutes: the Czech National Corpus Institute and the Institute of Theoretical and Computational Linguistics. In addition, more than two hundred external specialists from all over the Czech Republic are involved in collecting materials and coordinating partial activities.

3.1.4. *Corpus for the Tatar Language*

The *Tugan Tel Tatar Corpus* is a linguistic resource of the modern literary Tatar language. The project is implemented within the framework of the State Program "Preservation, Study and Development of the State Languages of the Republic of Tatarstan and Other Languages in the Republic of Tatarstan for 2014-2020". The corpus under development is addressed to a wide range of users: linguists, specialists in the field of Tatar, Turkic, and general linguistics, typologists, Tatar language teachers, cultural workers, and everyone who studies and is interested in the Tatar language. The volume of the corpus as of the end of 2018 is over 180 million word uses and contains texts of various genres (fiction, media texts, texts of official documents, educational literature, scientific publications, etc.).

3.1.5. *Corpus for the Kazakh Language*

The website of the *National Corpus of the Kazakh Language* has collected an electronic text fund for the Kazakh language. The volume of text in the corpus is 40 million. The fund of included texts has morphological, lexical, semantic, and phonetic-phonological designations. Research has been conducted on the compilation of the National Corpus of the Kazakh Language since 2009.

3.3. Instruments

3.3.1. *Corpus Instruments*

The central element of the study was a specialized corpus analyzer that allowed for frequency, concordance, and collocation analysis of texts. This tool provided the possibility of a detailed study of the use and context of mythonyms and theonyms in a large volume of texts.

3.3.2. *Linguistic Analysis Software*

Various linguistic analysis software programs were used, including syntactic and morphological structure analyzers. These tools allowed the automatic identification and analysis of grammatical and syntactic characteristics of mythonyms and theonyms in texts.

3.3.3. *Databases and Electronic Resources*

Electronic databases, including cultural encyclopedias, dictionaries, and folklore collections, were used to collect, store, and process information. This provided access to a wide range of materials for analysis.

3.2. Procedure

Data analysis was held in numerous stages. In the first stage, the technical correction of chosen texts occurred including the format standardization and unification of stylistic elements. Afterwards, the text passportization stage was completed which included the imposition of the metatextual markups for further cultural-semantic analysis. This process involved the identification and labeling of key text elements such as author, year of publication, genre, and type, as well as linguistic and extra-linguistic characteristics.

Methods of motivational analysis of cultural units were used to identify and interpret the cultural semantics of mythonyms and theonyms. The application of computer methods, including the development of a search and markup system, facilitated the process of quickly locating and analyzing the necessary data. These methods made it possible not only to collect and systematize a large amount of information but also to conduct an in-depth semantic and linguistic analysis providing qualitative results of the study.

The method of generalization of mythonyms and toponyms was used when creating a database of texts (folklore works, cultural encyclopedias, dictionaries, etc.) to be included in the corpus. This involved selection of texts consisting of mythonyms and toponyms, and using the method of selection, sorting, and technical correction. The collected texts went through a method of passportization entailing the imposition of metatextual markings to facilitate identification and analysis. Cultural semantics analysis played an important role involving the development of cultural semantic markings to comprehend the cultural semantics of mythonyms and theonyms. The approach encompassed methods of motivational analysis of cultural units using the "Mythonyms and Theonyms" methodology. Both linguistic and extralinguistic factors (metatextual, text author, year, text title, genre, type, etc.) were used for this analysis. Computer methods were also used to develop markup and create a "search" system that allows the user to quickly find the information they need. The rationality of linguistic analysis conducted on "real" language material was to obtain qualitative results and increase interest in corpus use (Chomsky, 2006).

Overall, the study was based on texts consisting of mythonyms and toponyms selected from various sources, including folklore works, cultural encyclopedias, and dictionaries. The data selection was based on the criteria which took into account their linguistic and cultural significance, and also, their representativeness in the world languages corpus. A particular focus was placed on texts that demonstrated the application of toponyms and mythonyms in many linguistic and cultural settings, as this offered a multifaceted perspective on how they operate in language.

4. Results

4.1. The Pragmatics of Theonyms and Mythonyms

Theonym pragmatics studies how people use and interact with these names, and what their sociocultural and linguistic aspects of the connected public practices and norms are. For instance, forms of greetings while addressing nature-deities may be included as a pragmatic aspect. Moreover, praise etiquette or rituals, the usage of these names in formal or informal contexts, also, translation features, and theonyms' interpretation could also be a part of it.

4.1.1. The German Language

The following features of German theonym pragmatics are a few specific examples:

- Address formats: when addressing divine creatures in religious situations, terms of address that convey awe and reverence may be employed. As an example, in Christianity, God is addressed with the form "Herr" (Lord), however, other forms of salutation, like "Herrgott" (Lord God) or "Allmächtiger" (Almighty), are also acceptable.
- Prayers and religious literature: German is a language that can be used in prayers and religious literature in the Christian religion. For example, the Lord's Prayer, a significant component of Christian tradition, is known in German as "Vater Unser" (Our Father).

- Proverbs and cultural expressions: Several German theonyms have religious origins and can be employed as metaphorical statements or proverbs. For instance, to “set heaven and hell in motion” (Himmel und Hölle in Bewegung setzen) is to use all necessary effort to accomplish a goal.
- Literature and art: German theonyms can be employed in literature and art to convey specific concepts or to set the mood. For instance, the Greek God Prometheus’ name is used in Friedrich Hölderlin’s poem Prometheus to represent man’s quest for freedom.

These are only a few illustrations of German theonym pragmatics. It is crucial to keep in mind, though, that these instances may vary based on the usage environment and the religious views of the person in question.

There are special phrases or expressions used to address deities or creatures when cursing or swearing in German. These are called theonyms and mythonyms of cursing. These are a few instances:

Theonyms for cursing:

- ‘Verdammt!’ A common way to swear is ‘curse!’ and it is occasionally taken as an oath. It conveys a pessimistic outlook or a desire to cause harm, although it is not connected to any particular deity.
- ‘Teufel!’ (Devil!) is a term for expletives or irritation. It is associated with the notion of the devil as a malevolent being.
- ‘Himmel, Arsch und Zwirn!’ One powerful example of a curse that incorporates the sky and conveys unpleasant feelings or a desire for retribution is ‘Heaven, Ass and Strings!’

Mythonyms for cursing:

- “Donnerwetter!” (Thunderstorm weather!) is a phrase that refers to thunderers but can also be used to convey surprise or irritation.
- “Zum Kuckuck!” (To the cuckoo!) is a swear phrase that is typically used in place of a swear word connected to the mythical or divine creature known as the cuckoo.
- “Verflixt und zugenäht!” (Cursed and sewed up!) is a curse term that denotes bad emotion or displeasure but lacks a closely associated deity.

These represent just some of the German theonyms and mythonyms for swearing. It is significant to remember that these expressions can differ in intensity and use depending on the situation and the individual.

4.1.2. *The English Language*

Curses and oaths can also be expressed in English using theonyms and mythonyms. Here are a few examples:

Theonyms for cursing:

- ‘God damn it!’ (God damn it!) is a profanity that is used in reference to God. It conveys irritation or dissatisfaction with a circumstance or occurrence.
- ‘Jesus Christ!’ can be used as a provocation or a sign of shock. It is connected to the identity of Jesus Christ.

Mythonyms for swearing:

- ‘Bloody hell!’ has a tendency to curse and is connected to hell. It conveys astonishment, displeasure, or outrage.
- ‘Damn it!’ is a term used to convey annoyance, unfavorable feelings, or displeasure. In this instance, the word ‘damn’ has a bad meaning.

Cultural mythonyms:

- ‘By Jove!’ (In the name of Jupiter!) refers to the Roman deity Jupiter. It expresses astonishment or shock.
- ‘By the beard of the Zeus!’ is an expression of astonishment or intense emotion that alludes to the Greek God Zeus.

These are only some of the English theonyms and mythonyms concerning swearing. It is crucial to keep in mind that these expressions might differ in power and usage depending on the situation and the individual. It is also important to remember that in some circumstances, including names from mythology and religion curses can be disrespectful or improper.

4.1.3. The Czech Language

Curses and vows can also be expressed with theonyms and mythonyms for cursing in Czech. Here are a few instances:

Theonyms for cursing:

- ‘Bože!’ A curse phrase that uses the name of God is ‘God’. It conveys astonishment, dissatisfaction, or disapproval of a circumstance or occurrence.
- ‘Ježíš!’ (Jesus!) can be used as a greeting of astonishment or as a curse. It is closely connected with the name of Jesus Christ.

Mythonyms for swearing:

- ‘Prokletí’ (Prokletí!) is a cursing term that conveys negative feelings, displeasure, or the want to curse rather than being connected to a particular deity or mythological creature.
- ‘Čert vzal!’ (The devil has stolen!) is a term of a curse connected to the devil. It’s employed to convey outrage or frustration.

Culture mythonyms:

- ‘Na Ráchtu!’ alludes to the mythical Czech figure Rácht, who is connected to disaster or bad luck. It is employed to communicate dissatisfaction or unpleasant feelings.
- ‘Kam s ním!’ (Where to with him!) is a term used to convey irritation or a bad attitude toward someone or something, although it has no clear connection to any legendary figure or deity.

These are just a few examples of theonyms and mythonyms of cursing in Czech. It is important to remember that these expressions can vary in usage and strength in different contexts and among different people. It is also worth noting that the use of religious and mythological names in curses may be offensive or inappropriate in some situations.

For centuries the Bible has been the most read and quoted book in England. There are so many biblical expressions and phrases that have entered the English language that it is a very difficult task to collect and list them. Here are the examples of biblicalism used in modern English (Gak, 1999):

- Apple of Sodom – a beautiful but rotten fruit; a deceptive achievement. The apple of Sodom is also used as an idiom meaning "deceptive appearance; beautiful but rotten fruit"
- The beam (mote) in one eye – the blind leading the blind
- The camel and the needle – an allusion to the Gospel: "It is easier for a camel to go through the eye of a needle than for a rich man to enter the kingdom of heaven"; Can a leopard change its spots?
- The crown of glory – the crown of glory, successful completion of any business, or process
- Daily bread – daily bread, means of subsistence; loaves and fishes – blessings of the earth (with bread and fish Christ fed hundreds of people who gathered to listen to him)
- The prodigal son – the prodigal son; the prodigal land – the land of life (British National Corpus)

Besides all proverbial sentences and expressions including various nominative (with the main word-noun), attributive, and adverbial phrases, many biblical verb phrases have entered English:

- To bear one's cross-carry one's cross–
- To condemn oneself out of one's mouth – to accuse oneself
- To escape by the skin of one's teeth – to barely escape, barely escape danger
- To kill the fatted calf – to welcome with open arms
- To laugh to scorn-precious laughter
- To sit under one vine and fig-tree – to sit under one vine and fig-tree (Oxford University Press, n.d.).

Besides phrases from the Bible, in English, as well as in other European languages, which are heirs of ancient culture, there are many proverbs, aphorisms, and figurative sayings, that originated with the ancient Greeks and Romans.

Thus, practically all languages present both common names and proper names of mythological and mythologized characters. Despite this, the peculiarities of personal names in mythology, the problems of their functioning, their connection with the name, and the reasons for the appearance of proper names in mythonymy are studied rather poorly. The works of Berezovich (2007), Vinogradova (2000), Levkievskaya (1999), Sklyarenko and Sklyarenko (2018) and others reflect some aspects of the above problem. As for the peculiarities of proper names in Turkic mythology, the ratio of proper and nominative names in its composition is practically not given.

Therefore, onomastic science cannot be limited at all to revealing the linguistic meaning of names, i.e. the language code. There are three different codes in the language. Because the content of an onomastic name contains such diverse information as 1) linguistic 2) cultural 3) historical codes (Uali, 2014). This is because linguistically diverse knowledge about different layers of language is often only available to the linguistic code, and surface information. Linguistic code is linguistic units, and names, in which information stores knowledge about the world. Cultural codes contain deep strata of non-linguistic information, reflecting the life experience and knowledge of the population about the environment since ancient times, traditional ethnic culture, beliefs, customs, and worldview.

4.1.4. *The Kazakh Language*

Mythonyms Bayanai and Bayanai have the same origin and have connections with bai (bai) of Turkic found in associations with various meanings such as "God, fate, contribution", "wealth, bai", and "man, master". The basis of bai (bai) is nomadic Kazakh Turkic languages and Bashkir which signifies "share, payment, fate", and "God of fate". This is reflected in expressions like bai kakal "got his share", bai kakal "wealth, symbols of abundance", bai koizh "ritual holiday", etc. (Khisamitdinova, 2010). The rich semantic base is recorded not only in Turkic but also in ancient languages. This is evidenced by materials such as the Ancient Turkic Dictionary (1969), Radlov (1899-1911), and the Etymological Dictionary of Turkic Languages (1974). Radloff's dictionary (1899-1911) features the mythonym Pazhan (in Teleut, Tuba, and Lebedian languages), where the lexeme's meaning corresponds to "good spirit" or "one of the highest gods". In addition, the Uyghur language records the mythonym "God", which also seems to have the base baj (bai) / paj (pai) (Radlov, 1899-1911). Radlov's dictionary (1899-1911) depicts Kyrgyz lexemes, where "rich" denotes "wizard" and "wealth" means "lie".

Thus, modern onomastic research is aimed at defining the structural onomastic space, which summarizes the ethno-cultural experience, peculiar to the bearers of a certain national culture. Names with special associative potential have become a well-studied area in corpus linguistics as signs of cultural tradition due to their increased textual significance.

4.1.5. *The Tatar Language*

The study of the representativeness of mythonyms and theonyms in the Tatar language text corpus is aimed at analyzing the internal structures of the language related to mythological and religious elements. Analyzing the Tatar language text corpus is planned to reveal the distribution and

contextual use of mythonyms (mythological terms or images) and theonyms (religious terms and names of deities) in language practice. Examples of the use of mythonyms and theonyms in Tatar language texts may include:

Mythonyms (mythological images):

- "Irek was as brave as Tulpar, a legendary horse from Tatar mythology". In this case, "Tulpar" represents a mythological horse used to compare to a brave and courageous character.

Theonyms (religious terms):

- "We pray to Allah for the wisdom and guidance of our people". In this example, "Allah" is a theonym referring to God in Islam and is used in the context of prayer and blessing.

Combined usage:

- "In telling legends about Bai-Batyr, people marvel at his strength, comparing him to a demigod". In this case, "Bai-Batyr" is a mythological hero in the Tatar epic, and the sentence draws a parallel with the concept of a demigod to emphasize his incredible strength.

These examples show how mythonyms and theonyms occur in Tatar language texts, demonstrating their role in linguistic culture and expression.

4.2. Theonyms and Mythonyms in the Corpus of Texts

4.2.1. The German Language Corpus

In a corpus of German texts, theonyms can be used in a variety of contexts, including religious texts, literature, folklore, etc. Some of the most common theonyms in German include:

Gott – "God". This theonym is used in the Christian tradition as well as in some other religious traditions.

God – Powerful supernatural beings are the most important characters in developed religious mythologies. The image of gods merges the features of cultural heroes-demiurges, patrons of initiation rites, and various spirits. Creative functions are combined with the control of individual forces of nature and the cosmos as a whole, and the management of the life of nature and mankind. The concept of the supreme deity of the polytheistic pantheon evolves in the higher religions to the monotheistic image of a single god-creator and ruler of the universe. This concept is reflected in the screenshot below (Figure 1).

Figure 1

The Theonym Gott, "God" in the German Corpus

The screenshot shows the 'corpuseye' search interface. At the top, there are navigation links: 'Help', 'Grammatical information', and 'Taglist'. The search results are displayed in a table with columns for 'Field', 'By', 'Offset', 'Freq items', 'Collapse case', 'Collapse diacritics', and 'Export'. The search criteria are: 'word="Gott" %c' in the 'GER_EUROPARL' corpus, with 257 results found. The results are sorted by 'Word' and 'freq'. The first few results are:

Field	By	Offset	Freq items	Collapse case	Collapse diacritics	Export
Word	Left Context	0	100	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Word	Right Context	0	100	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Word	Left Edge	0	100	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Word	Right Edge	0	100	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Word	Dep Head	0	100	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

The search results are displayed in a list of text snippets, each with an 'Export' button and an 'INFO' link. The snippets are:

- Logik - fast hätte ich gesagt von einem einzigen **Gott** - beherrscht , nämlich der Logik des arbeitenden Geldes
- Lage in der russischen Öffentlichkeit . Es gibt **Gott** sei Dank Kritik . Es gibt die Kritik der Soldatenmütter
- Standpunkt des Rates abgeschwächt wurde , es aber **Gott** sei Dank den großen Fraktionen hier nicht gelungen
- . Wir dürfen Afrika , den nicht nur von **Gott** , sondern auch von Europa völlig allein gelassenen
- die wir alle in unseren Territorien haben und die **Gott** sei Dank eine Minderheit darstellen , die Fackel der
- und bedauerliche Zwischenfälle , aber was es **Gott** sei Dank nicht mehr gibt , ist die Massenvertreibung
- auch einbeziehen müssen , denn sollten wir , was **Gott** verhüten möge , eine Krise haben , werden alle Truppen
- erste Kriegstreiber , der dies schwört . Möge uns **Gott** vor den schrecklichen Folgen der großsprecherischen
- Euro sein , und das wäre in diesem Fall weiß **Gott** zu wenig . Dann ändert sich nichts . Es entsteht kein
- wirklich zu verstehen , wenn der Blick auf **Gott** , den Schöpfer des Universums und also auch des Menschen
- und sie durch ein individualistisches Magma ohne **Gott** und Herren zu ersetzen . In dieser Konzeption sind
- Frage an Sie , Herr Kommissar : = hat bis jetzt **Gott** sei Dank keinen einzigen Fall von BSE gehabt . Werden
- auszusponieren . Wohin führt uns das ? **Gott** , schütze uns vor unseren Freunden . Um die Feinde

Allah – "Allah". This theonym is used in the Islamic tradition. In ancient Arabian mythology, the supreme deity worshipped in Northern and Central Arabia was the god-predecessor and demiurge. Allah is both the god of the sky and rain. Probably, in accordance with the Semitic tradition, the word "Allah" replaces the name of God. Giving the word the article implies that the word is referring to a well-known god. Originally one of the main deities of the Arabian pantheon, but not associated with any particular ethnic group. This concept is reflected in the screenshot below (Figure 2).

Figure 2

The Theonym Allah, "Allah" in the German Corpus

The screenshot shows the 'corpuseye' search interface. The search criteria are: Field: Word, By: Word, Offset: 0, Freq items: 100. The search results are displayed in a table with columns for 'Export' and 'INFO'. The search results are as follows:

Export	INFO
Export	INFO
Export	INFO
Export	INFO
Export	INFO
Export	INFO
Export	INFO
Export	INFO

Engel – "Angel". This theonym is used in the Christian tradition and in other religious traditions. An angel is a spiritual, incorporeal being who communicates the will of God and possesses supernatural powers. Traditionally, an angel is depicted as an anthropomorphic being with wings behind his back. More broadly speaking, angels are benevolent spiritual entities. This concept is reflected in the screenshot below (Figure 3).

Figure 3

The Theonym Engel, "Angel" in the German Corpus

The screenshot shows the 'corpuseye' search interface. The search criteria are: Field: Word, By: Word, Offset: 0, Freq items: 100. The search results are displayed in a table with columns for 'Export' and 'INFO'. The search results are as follows:

Export	INFO
Export	INFO
Export	INFO
Export	INFO
Export	INFO
Export	INFO
Export	INFO
Export	INFO
Export	INFO
Export	INFO
Export	INFO
Export	INFO
Export	INFO
Export	INFO
Export	INFO
Export	INFO
Export	INFO
Export	INFO
Export	INFO
Export	INFO
Export	INFO

Zeus – "Zeus". This theonym is used in Greek mythology. Zeus is the god of the third generation. According to Greek mythology, at the beginning of time, the gods fought each other for the right to rule over the Earth. Kronos (Zeus' father), the youngest of the children of Uranus (Heaven) and Gaea

(Earth), overthrew his father and became the supreme deity. This concept is reflected in the screenshot below (Figure 4).

Figure 4

The Theonym Zeus, "Zeus" in the German Corpus

corpuseye [Help](#) | [Grammatical information](#) | [Taglist](#)

sort freq rel

Field: Word

Left Context
Right Context

By: Left Edge
Right Edge
Dep Head

Offset: 0

Freq items: 100

Collapse case

Collapse diacritics

Export

Field: Word

Max rows: 300

Count n-grams
Export n-grams

[Refine search](#)

[New search](#)

Searched for: [word="zeus" %c]
In corpus: GER_EUROPARL
Found 2 results (2).

1

Export All

Export INFO

Export INFO

Poseidon . Europa war die Geliebte des Zeus und Mutter des Minos , des Sarpedon sowie des Rhadamanthys

Elemente . Europa war die Geliebte des Zeus und die Mutter von Minos , Sarpedon und Rhadamanthys

In the corpus of German texts, theonyms can be used both to denote (label) the name of a deity and to denote (label) related concepts and notions. A study of the use of theonyms in German can help in understanding the religious, mythological, and cultural representations associated with these theonyms in German and German culture in general.

Analyzing theonym usage in German is an intriguing way to undertake linguistic study. Firstly, theonyms are frequently used as nouns. The phrase "Ich glaube a Gott" (I believe in God) is one example of how the word "Gott" (God) may be employed in many different contexts. "Gott ist unglaublich mächtig" God is incredibly powerful or "Gott kommt yu Ihnen" God comes to you or "May God grant you blessing. Secondly, proper names can be made out of theonyms. The name "Maria" (Maria), for instance, can be used in the following situations: "Maria ist ein häufiger name" (Maria is a name that is fairly common). "Die Kirche hat eine Statue von Maria" (The church has a statue of Maria.). "Maria hat mir geholfen" (Maria helped me). Thirdly, theonyms can be used as adjectives. For example, the word "heilig" (saint) can be used in the following contexts (Table 1):

Table 1

The Word "Heilig" (Saint)

Heiliger Geist (Holy Spirit)	a term used in Christian theology to describe the third person of the Trinity.
Heilige Stadt (Holy City)	a term used to describe cities that are considered sacred because of their religious history or significance.
Heiliger Krieg (Holy War)	a term used to describe military conflicts that are conducted in the name of religion or in defense of religious ideals.
Heiliges Land (Holy Land)	a term used to describe a territory that has religious significance, such as Israel for Judaism, Christianity, and Islam.
Heilige Familie (Holy Family)	a term used to describe Jesus Christ, Mary, and Joseph, who are considered saints in the Christian tradition.
Heiliger Ort (Holy place)	a term used to describe a place that has special religious significance or is associated with a sacred event or legend.
Heiliges Kreuz (Holy Cross)	a term used to describe the cross on which, according to Christian belief, Jesus Christ was crucified.

Heilige Jungfrau (Holy Virgin)	a term used to describe Mary, the mother of Jesus Christ, who in Christianity
Heiliger Berg (Holy Mountain)	a term used to describe mountains that are considered sacred because of their religious history or significance.
Heiliger See (Holy Lake)	a term used to describe lakes that are considered holy because of their religious history or significance, such as Lake Tiberias in Israel for Judaism and Christianity.
Heiliger Schutzpatron (Patron Saint)	a term used to describe holy people who are considered patrons of certain groups of people, professions, or geographic areas.
Heiliges Feuer (Holy fire)	a term used in various religious rituals that symbolizes holiness and purification.
Heiliges Opfer (Holy sacrifice)	a term used in religious rituals that symbolizes the offering of some sacrifice or offering to God as a sign of gratitude or repentance.

In conclusion, theonyms, that is, words related to the name of a deity, can be used not only as names of deities themselves, but also as adjectives describing sacred places, objects, rituals, and traditions. This shows the importance of religion in people's culture and language, and how it influences our vocabulary and cultural practices.

A corpus analysis of the ancient mythology of the Scandinavian world provides an opportunity for reflection in relation to modern Scandinavian society. The study of the relationship between the supreme gods and the representatives of the ancient dynasties of the mythological world helps in the study of the modern mentality and identity of Norwegians, Swedes, and Danes. The benefit of linguistic analysis and actualization of knowledge about the traditions, beliefs, and habits of the Scandinavian public also helps in the issue of using international policy instruments such as the soft power of the state, which undoubtedly still plays a crucial role in building the image and attractiveness of the Nordic Scandinavian states of Northern Europe (Nye, 2004).

4.2.2. The English Language Corpus

There are many mythonyms in the English language that originate from ancient Greek, ancient Roman, Scandinavian, Egyptian, and other mythologies. Table 2 encapsulates some of the most common mythonyms in English.

Table 2

Some of the Most Common Mythonyms in English

Herculean task	Herculean labor; labor that requires special effort and endurance.
Achilles' heel	Achilles' heel, a weakness that can lead to defeat or failure.
Pandora's box	Pandora's box, something that could lead to undesirable consequences or problems.
Trojan horse	Trojan horse, a stealthy method of infiltration or attack.
Siren song	siren song, an attractive but dangerous proposition.
Midas touch	Midas touch, the ability to increase wealth.
Phoenix rising from the ashes	a phoenix rising from the ashes, a symbol of recovery or rebirth after disaster.
Cupid's arrow	Cupid's arrow, symbol of love and passion.

Studying the use of mythonyms in a corpus of English texts can help in understanding the cultural and literary references that authors may use in texts. It can also help in the study of language and literature, and in analyzing the use of mythology in modern society. Mythonyms in English can be used in various fields such as literature, religion, science, art, politics, and even sports. For example, mythonyms are often used in newspaper headlines to draw attention to news stories and create emotional coloring. Mythonyms can also be used to create new words and expressions.

In addition, mythonyms can be used to create identification and social connections between people who share a common knowledge of mythology. For example, using mythonyms in conversation can help establish rapport with other people who share an interest in mythology or the culture on which it

is based. In today's culture, mythonyms are often used in the names of companies, brands, products, and services. For example, the company *Nike* is named after the ancient Greek goddess of victory, Nike (see Figure 5), and *Amazon* is named after the ancient Greek mythology of the Amazons, a female warrior (see Figure 6).

Figure 5

The Word Nike in British National Corpus



The screenshot shows the 'corpuseye' search interface. The search criteria are: Field: Word, By: Right Context, Offset: 0, Freq items: 100. The search results are as follows:

Field	Word	Count
Word	Nike	2

Search results (2):

- Union to spread throughout the world , Adidas , **Nike** and other sponsors who want to see results , transfers
- goods companies in Europe - Adidas , Reebok and **Nike** - who are now ensuring that there are fair wages from

The word Nike occurs 2 times in the corpus of English texts. In the above screenshot, the word Nike is written in the corpus text as a known modern brand name. Hence, the name of the god Nike in Greek mythology was not included in the above corpus of texts.

Figure 6

The Word Amazon in British National Corpus



The screenshot shows the 'corpuseye' search interface. The search criteria are: Field: Word, By: Right Context, Offset: 0, Freq items: 100. The search results are as follows:

Field	Word	Count
Word	Amazon	14

Search results (14):

- a project that authorises the use of 50% of the **Amazon** rain forest , which we know represents a huge genetic
- who bought this book also bought ? ' With **amazon** you can go on to make and accumulate purchases , so
- believe that a standardised portal along **amazon** lines - partly , but not exclusively , computer-driven
- of European Socialists , on the mining in the **Amazon** (Venezuela) ; -B4-1141/96 by Mr Eisma and others
- and Reformist Party , on the mining in the **Amazon** (Venezuela) ; -B4-1162/96 by Mr Kreissl-Dörfler and
- in the European Parliament , on mining in the **Amazon** ; Bulgaria
- 1991 . It covers an area of 84 , 000 km2 in the **Amazon** state and comprises , in addition to the national parks
- In this **Amazon** state mining deposits are worked , especially by gold
- the European Parliament , on forest fires in the **Amazon** region of Brazil ; -B4-0886/97 by Mrs González Álvarez
- Nordic Green Left , on the fires in the Brazilian **Amazon** region ; -B4-0845/97 by Mrs García Arias and others
- that the forests are on fire in the Brazilian **Amazon** . For many years human hands have been making irreparable
- of inspiring the Brazilian authorities to an **Amazon** development strategy which is better adapted to environmental
- or agreement involving all the countries of the **Amazon** , under the auspices of the United Nations , clearly
- cost price to raise cattle on the ashes of the **Amazon** rainforest , it would be catastrophic for the environment

The word Amazon occurs 14 times in the corpus of texts in English. Of these, it is written 2 times in two sentences in connection with the current Amazon website. In specific, 10 times as an onima, and 2 times in a text about Amazon warriors in mythology.

Mythonyms are also used in cultural industries such as movies, television, music, and video games. For example, *the God of War* series of games is based on Greek mythology and includes characters such as Zeus, Poseidon, and Hercules (Figure 7).

Figure 7
The Phrase God of War in British National Corpus

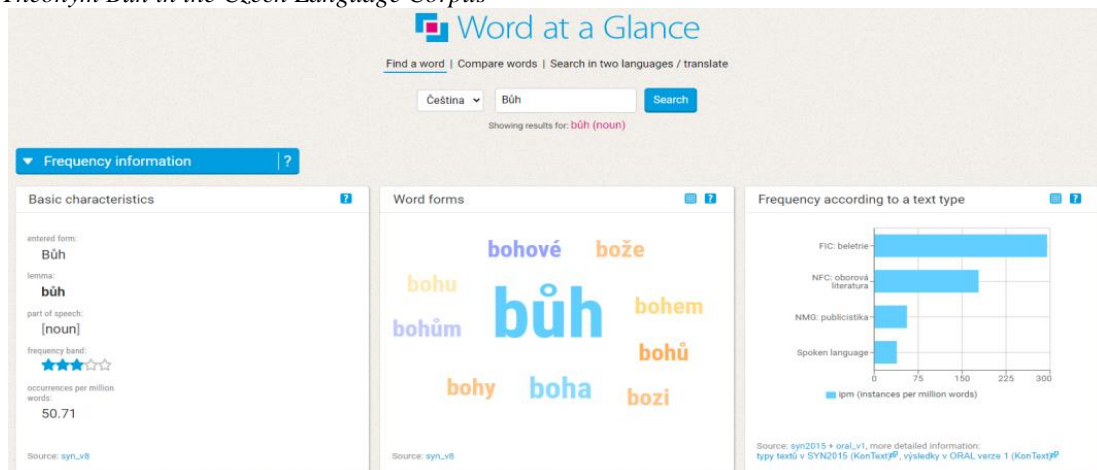


God of War emerges in game form. In the English corpus of texts, God of War is embedded as a constant phrase in the essence of what is happening in the modern European Union. From this we can clearly see that mythology today influences the culture of language expression and thought.

4.2.3. The Czech Language Corpus

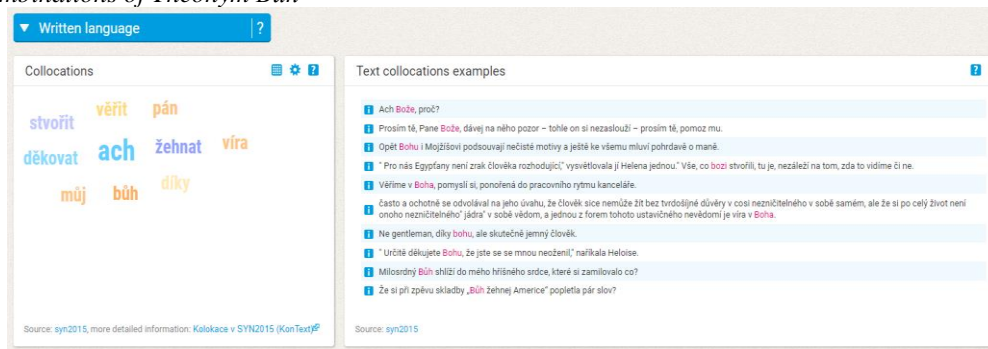
The Czech language has the grammatical markup *Bůh*, i.e., theonymy of God. The 9 word forms of theonymy *Bůh* are also shown. The frequency of occurrence in relation to 4 text types is calculated. The frequency of occurrence of *Bůh* theonymy is 12,279 in fiction, 7,020 in regional literature, 2,220 in journalism, and 246 in colloquial speech (Figure 8).

Figure 8
The Theonym *Bůh* in the Czech Language Corpus



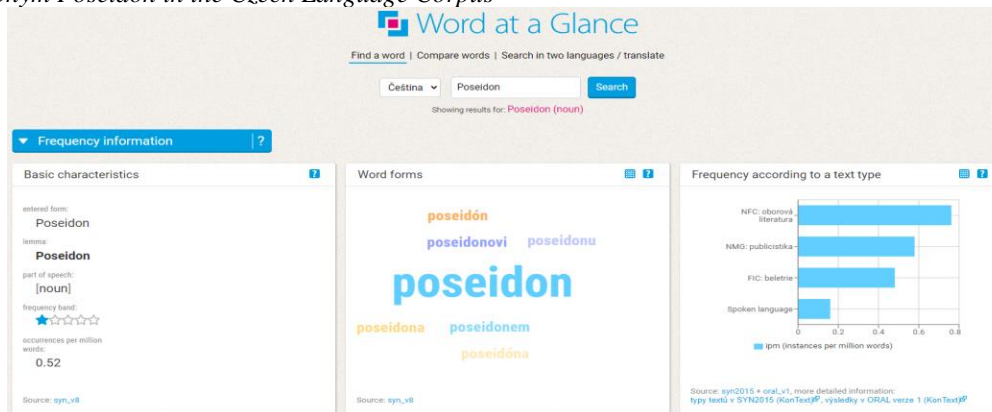
In written language, collocations with *díky* (thank you), *věřit* (believe), *děkovat* (thank), *víra* (faith), *pán* (Lord), and *stvořit* (create) are frequent. The screenshot below (Figure 9) shows examples of the sentences.

Figure 9
Word Combinations of Theonym Bůh



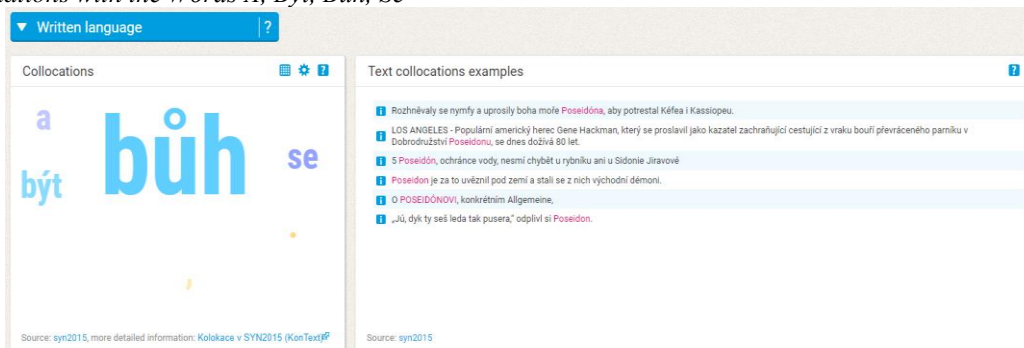
The theonym *Poseidon* (God of the sea) has a grammatical markup. In Figure 10, 6 word forms of theonym are shown. The frequency of theonym *Poseidon* in relation to 4 types of text is in regional literature: 30, in journalism: 23, in fiction: 20, and in colloquial speech: 1.

Figure 10
The Theonym *Poseidon* in the Czech Language Corpus



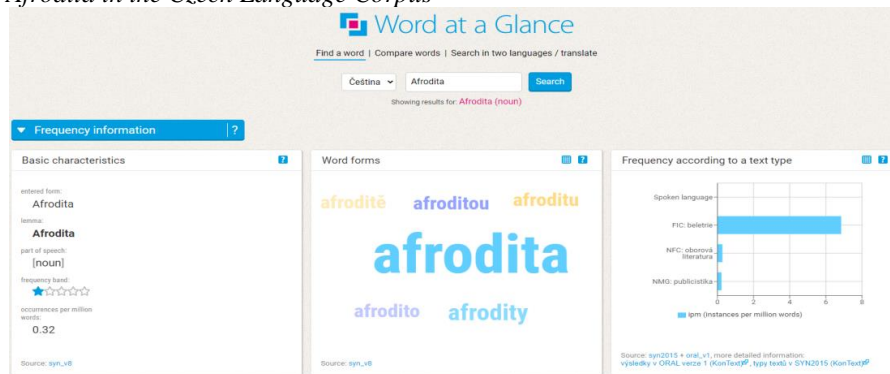
The written language has a high frequency of combinations with the words *a* (and), *být* (to be), *bůh* (God), and *se* (with). The screenshot below (Figure 11) shows examples of the sentences.

Figure 11
Combinations with the Words *A*, *Být*, *Bůh*, *Se*



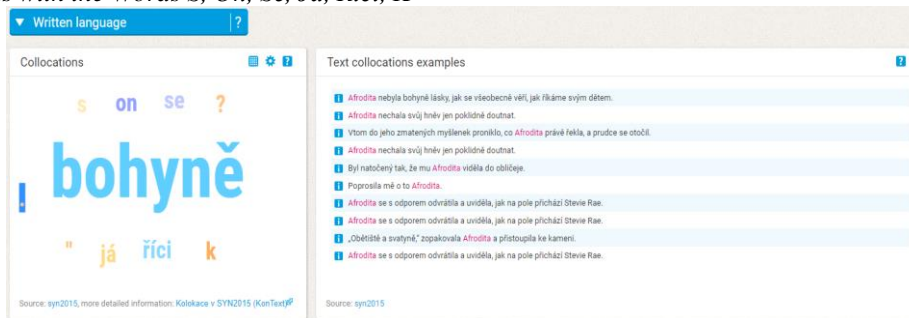
The theonym *Afrodita* (Aphrodite) has a grammatical markup in the Czech Corpus. In Figure 12, 5 word forms of theonym are shown. The frequency of theonym *Afrodita* in relation to 4 text types is: in regional literature: 11, in journalism: 9, in fiction: 285, and in colloquial speech: 0.

Figure 12
The Theonym Afrodita in the Czech Language Corpus



In written language, there is a high frequency of utilizing the words s, on, se, já, říci, k., and dot (.), along with a notable tendency to combine quotation marks (" ") with punctuation marks. Figure 13 shows the words in the sentence.

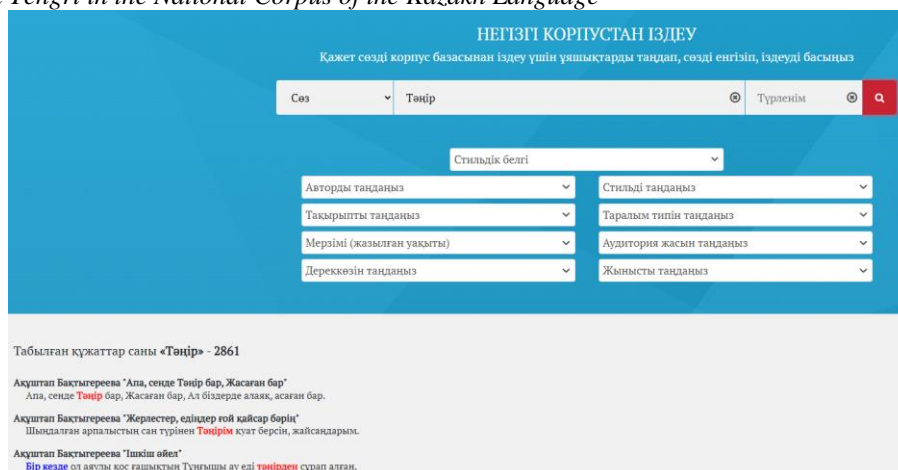
Figure 13
Combinations with the Words S, On, Se, Já, Říci, K



4.2.5. The Kazakh Language Corpus

Research has shown that before the spread of Islam in Kazakhstan, Turkic peoples believed in numerous deities such as the moon, sun, sky, earth, God, spirits, and ancestors (Kulmanova et al., 2022). According to the search results in the National Corpus of the Kazakh language theonyms, Tengri (Master of the blue sky), Allah, God, Creator, and Heavenly, 2861 documents were found, including the total number of results for the styles fiction, science, speech, and business (Figure 14). The search result for the query "business styles" was 0.

Figure 14
The Theonym Tengri in the National Corpus of the Kazakh Language



The conclusion that can be drawn from this is that the color "blue" is undeniably popular among Turkic-speaking peoples and is one of the colors representing national culture and national worldview due to its cognitive nature (Pansat & Khalikova, 2023). The mythological worldview of the Kazakh people (Tengrianism) and religious notions are reflected in phraseology, proverbs, winged words, and folklore texts (fairy tale, mythical story, legend, afsana, kisa). For example, the Kazakh concept and existence of the word Kok (blue) has a deep meaning and is perceived as a symbol of God. Therefore, Kazakhs, raising their hands to the sky, ask May God bless you!, May God fulfill your wishes!, May God punish you!, ascended to heaven – to God, raised to heaven – to God; etc. (Kenesbaev, 1977).

In general, the worship of Turkic-speaking peoples is the sky, the God of heaven. The word "Tanir" means "God, Allah, creator". In the "myth of the creation of Tengri", the creation of Tengri from the ocean is described as follows: "Long ago there was one boundless ocean, in which there was neither heaven nor earth. Once in the middle of the ocean appeared a white luminous golden egg, from which the future progenitor of the whole world Tengri was created (Kazakh folklore in the collection of Potanin, 1972). Phraseological units are units resulting from expressive indirect conception created with the help of certain signs (Bekeyeva et al., 2021). In phraseological expressions with the semantics of ill-wishing Tanirden tap (Find from the God), Tanir atkyr (may God punish!), Tanirge ne zhazdym (what is my fault?!), Tanir zharylkasyn (may God bless!), Tanir isi (God's Providence), and Tanir zhazsa (as God wills) also occur in cultic meaningful phrases, such as God blessed (Kenesbaev, 1977).

Currently, the subcorpus of the theonyms and mythonyms is being developed in the National Corpus of the Kazakh language. The subcorpus of mythonyms is subdivided into the themes mytho-anthroponym, mythotoponym, mythozoonyme, and mythophytonym. Each mythonymic name has cultural semantics. And the subcorpus of theonyms consists of names of creators, prophets, names of angels, and companions (Table 3).

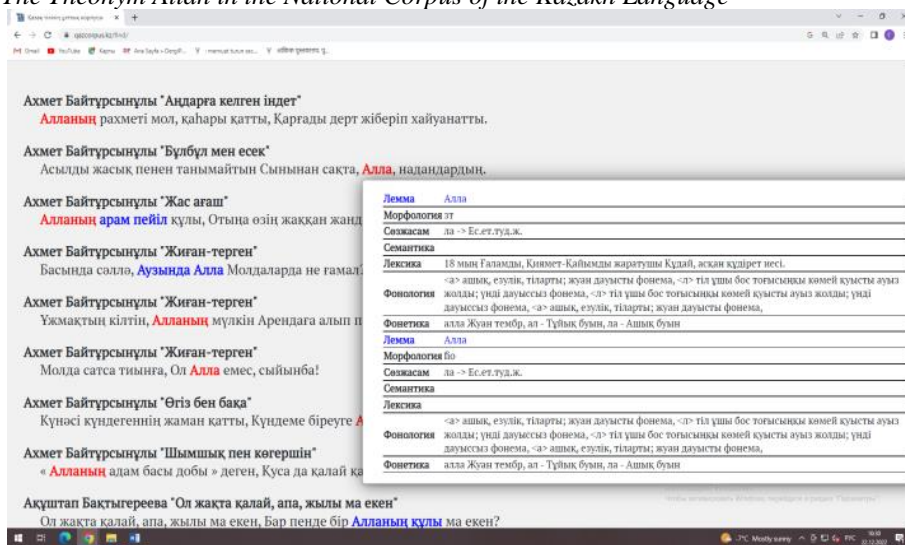
Table 3

Subcorpus of Mythonyms and Theonyms in the National Corpus of the Kazakh Language

SUBCORPUS OF MYTHONYMS		THEONYMS SUBCORPUS	
Mythoanthroponym	Pandemonium Devil, Giant, Devourer, Copper, Fairy, Devil, Angel, etc.	Creator	Allah, Tengri, God, etc.
Mythoponym	Sorkuduk, Mount Arafat, fiery Darius, etc.	Prophets	Mohammed, Musa, Jesus, etc.
Mythozoonyme	Dragon, Samruk, etc.	Angels	Azreil, Gabriel etc.
Mythophytonym	Baiterek, Darak, Zheken, etc.	Companions	Omar, Abu Bakr, Osman, Ali and others.

After the full introduction of theonyms into the corpus as markup, the cell with markup related to the above cultural-semantic meaning comes out, and stable word combinations are replenished (Figure 15). Also, proverbs and sayings, winged words, and ethnographic works will be included in the base of the National Corpus of the Kazakh language. A number of synonyms of theonym Allah/Allah are not specified in the base of the modern corpus. When introducing the markings of theonyms into the corpus, an algorithm for coding a number of synonyms of nominal and other mythonyms and theonyms is provided. For example, the synonymic series of the word Allah is formed by the words Zharatushy, Zharatkan, Tanir, Kudai, Zhappar, Rabby, and Tagala with the general meaning "Creator". The 99 artistic names of Allah are also included as markings, and each name is revealed with a breakdown by breadth and meaning (Figure 15).

Figure 15
Reflection of The Theonym Allah in the National Corpus of the Kazakh Language



By structure, the personal names of Turkic mythological characters can be divided into three groups. The first group includes names consisting only of nomenclatural terms: tengri 'god, deity', ata 'elder, father, grandfather', baba 'father, ancestor, grandfather, elder', ana/ine 'mother', e)е/ue/eje 'master spirit', and batyr 'bogатыr, hero'. The nomenclatural term Tengri is often found among the personal names of heroes of Turkic mythology. The ancient Turkic mythonym Tengri is reflected in the dictionary of Mahmud Kashgari, monuments of ancient Turkic writing, etc. (Sklyarenko & Sklyarenko, 2018). In various phonetic variants, it is represented in many modern Turkic languages; for example, the Bashkir proverb Kurshe hsgky – Tetsre hsgky "The right of the neighbor is like the right of Tengri".

Figure 16
Occurrence of Theonym Allah in Phraseological Expressions

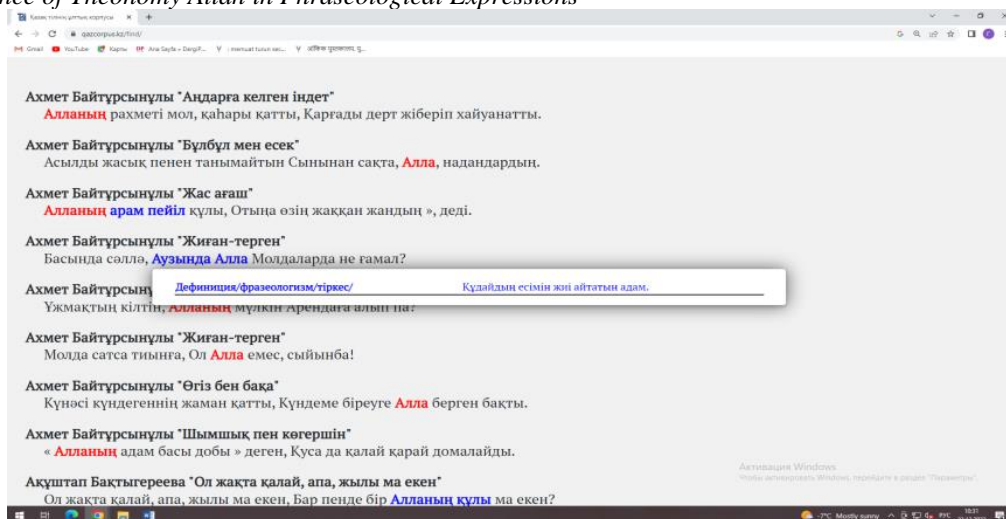


Figure 16 is a screenshot image taken from the Kazakh language corpus database. Only the definition of the phraseologism is given here, but it is not complete. The base of the modern corpus includes works of fiction, scientific texts, examples from newspapers and magazines on the Internet site, and various business texts. Then mythonyms and theonyms were identified. And in order for mythonyms and theonyms to be identified and distinguished from other texts, we supplement the text base with

legends, fairy tales, mythical stories, and songs reflected in the consciousness of the population. Only then, on the basis of newly collected texts of mythonyms and theonyms, the cultural-semantic meaning is revealed comprehensively and completely.

4.2.6. The Tatar Language Corpus

In the Tugan Tel corpus of the Tatar language, the frequency of the theonym Paigambar (Prophet), out of 347 documents found in the corpus, is 677 (Figure 17). A number of 665 phrases were produced in the context search. The number of words is 1. Hence, in the Tugan Tel corpus, the theonym Prophet is found in sentences used only in the root. From this, we noticed that in the Tugan Tel corpus, the morphological composition (markup) of the theonym Prophet was specified in relation to only one lemma (Tatar language corpus).

Figure 17
The Theonym Paigambar in the Tatar Corpus

The screenshot shows the search interface for the Tugan Tel corpus. The search term is 'пайгамбар' (paigambar). The results table is as follows:

#	словоформа	лемма	морфологические свойства	Пример контекста	Количество результатов
1.	пайгамбар	пайгамбар	N+Sg+Nom	Яджарь курсетуга сирак куела, Темэн елкасеннан читка чыкканы юк диярлек истанбулның "хирка шериф" мечетенде пайгамбар галейниссаламның хиркасы – эс киеме саклана. // Источник: http://kiziltan.ru/	677

The frequency of the theonym Rasul (one of the beautiful names of Allah), obtained from 164 documents found in the corpus, is 277 (Figure 18). A number of 273 phrases were found in the context search. The number of collocations is 1. Hence, the Tugan Tel corpus contains the theonym Rasul in sentences used only in the root. From this, we noticed that the morphological composition (marking) of the theonym Rasul with respect to only one lemma is specified in the Tugan Tel corpus.

Figure 18
The Theonym Rasul in the Tatar Corpus

The screenshot shows the search interface for the Tugan Tel corpus. The search term is 'расул' (rasul). The results table is as follows:

#	словоформа	лемма	морфологические свойства	Пример контекста	Количество результатов
1.	расул	расул	N+Sg+Nom	хазрати Расул акрам Аллаһы Тагаланең иң якын илчесе буларак, тугры юлны дөнья һәм ахирәттә мәңгелек бәхет-сагадатка, рахмәтка ирешү юлын Ислам дине белән курсатте. // Источник: http://kiziltan.ru/	277

The frequency of the theonym Hoda, out of 209 documents found in the corpus, is 437. A number of 434 phrases were produced in the context search. The number of word forms is 1. Hence, in the Tugan Tel corpus, the theonym Hodai is found in sentences used only in the root. From this, we noticed that the corpus of Tugan country specifies the morphological composition (markup) of the Theonym Hoda with respect to only one lemma (Tatar language corpus).

The study of mythonyms and theonyms in different languages and cultures has revealed a profound relationship between language, religion, and culture. For example, in English, phrases such as "Achilles' heel" or "Pandora's box" carry a rich mythological history that influences contemporary cultural perceptions. These images echo similar concepts in other cultures, emphasizing the universality and versatility of mythological language.

5. Discussion

The goal of our study was to undertake a comparative investigation of the use of mythonyms and theonyms in language corpora from diverse world languages, including English, German, Czech, Kazakh and Tatar. We attempted to determine the frequency with which these terms were used in various text styles, as well as to assess the cultural code linked with these names.

The analysis of mythonyms and theonyms in onomastic subcorpora found interesting patterns. For example, in English, we have a high number of references to theonyms in literary writings, which could imply a profound cultural saturation of religious conceptions in the language's literature. In the context of the German language, mythonyms receive special emphasis in scientific publications, presumably due to academic tradition and the inclusion of mythical elements in research. The cultural traits of each language provide insight into the causes of the discovered patterns. For example, the Czech language makes extensive use of theonyms in phraseological units that may be identified with the country's distinct cultural and historical fabric.

The results further highlight the necessity of studying the pragmatics and cultural semantics of theonyms. In the words of the study, the use of theonyms in phraseological units can be interpreted as a way of preserving and transmitting cultural values. Furthermore, the study has an influence on our knowledge of language and culture settings, and it motivates future research. It is especially important to continue investigating the influence of mythonyms and theonyms on the formation of collective consciousness in diverse communities. Furthermore, the link between the use of these phrases and the dynamics of language development in the context of globalization requires additional investigation. Thus, our research opens up new avenues for understanding language characteristics and cultural codes, and more research in this area can broaden our knowledge and shed light on a variety of remarkable linguistic and cultural occurrences.

Disclosure Statement

The authors claim no conflict of interest.

Funding

The research paper was obtained within the framework of the program BR 21882249 "Improvement and Expansion of the Subcorpus of the National Corpus of the Kazakh Language as a Means of Intercultural Communication".

References

- Baker, H., Gregory, I., Hartmann, D., & McEnery, T. (2019). Applying geographical information systems to researching historical corpora. In M. Mahlberg & V. Wiegand, (Eds.), *Corpus linguistics, context, and culture* (pp. 109-136). Walter de Gruyter. <https://doi.org/10.1515/9783110489071-005>
- Baker, P. (Ed.). (2012). *Contemporary corpus linguistics* (1st ed.). Continuum.

- Bekeyeva, N., Bissengali, A., Mankeyeva, Z., & Nurdauletova, B. (2021). Phraseological expressions in the Turkic language: Comparative analysis. *International Journal of Society, Culture & Language*, 9(2), 29-40.
- Berezovich, E. L. (2007). *Language and traditional culture: Ethnolinguistic studies*. Indrik.
- Biber, D. (1993). Representativeness in corpus design. *Literary and Linguistic Computing*, 8(4), 243-257.
- British national corpus. (2022). *British National Corpus online*. <http://www.natcorp.ox.ac.uk/>
- Chomsky, N. (2006). *Language and mind* (3rd ed.). Cambridge University Press. <https://doi.org/10.1017/CBO9780511791222>
- Czech language corpus. (2023). *Word at a Glance*. <https://www.korpus.cz/slovo-v-kostce/search/cs/B% C5% AFh>
- Dang, D., & Vartiainen, T. (2019). Digital strategy patterns in information systems research. In D. Xu, J. Jiang, H. W. Kim (Eds.), *Proceedings of 23rd Pacific Asia Conference on Information Systems* (pp. 1-18). PACIS.
- Danish language corpus. (2023). *KorpusDK*. <https://ordnet.dk/korpusdk/teksteksempler/sogeresultat>
- Gak, V. G. (1999). *Phraseology in the context of culture*. Institute of Linguistics.
- German Language Corpus. (2023). *German corpus*. <https://corp.hum.sdu.dk/cqp.de.html>
- Kenesbaev, I. K. (1977). *Phraseological dictionary of the Kazakh language*. Gylym.
- Khisamitdinova, F. G. (2010). *Mythological dictionary of the Bashkir language*. Science.
- Krasina, E. A., & Perfilieva, N. V. (2015). *Fundamentals of philology: Linguistic paradigms*. Science.
- Kulmanova, Z. B., Zhirenov, S. A., Mashinbayeva, G. A., Orynbayeva, D. G., Abitova, Z. S., Babayeva, K. S. (2022). Reflection of the religious worldview in language. *International Journal of Society, Culture & Language*, 10(3), 37-38. <https://doi.org/10.22034/ijscsl.2022.544382.2484>
- Levkievskaya, E. E. (1999). Mythological character: The correlation of name and image. In E. E. Levkievskaya (Ed.), *Slavic etudes: Collection for the anniversary of S. M. Tolstoy* (pp. 243-257). Indrik.
- McEnery, T., & Wilson, A. (2001). *Corpus linguistics* (2nd ed.). Edinburgh University Press.
- McEnery, T., Xiao, R., & Tono, Y. (2006). Corpus-based language studies. *An advanced resource book*. Routledge.
- Motschenbacher, H. (2020). Corpus linguistic onomastics: A plea for a corpus-based investigation of names. *NAMES a Journal of Onomastics*, 68(2), 88–103. <https://doi.org/10.1080/00277738.2020.1731240>
- Nadeliev, M., Nasilov, D. M., Tenishev, E. R., & Shcherbak, A. M. (Eds.). (1969). *ATD – Ancient Turkic Dictionary*. Nauka.
- Nye, J. S. (2004). *Soft power: The means to success in world politics*. Public Affairs.
- Oxford University Press. (n.d.). Biblical verb phrases. Retrieved January 10, 2023, from <https://en.oxforddictionaries.com/>
- Pansat, Z., & Khalikova, N. (2023). Semantic features of color in emotional, expressive words: The concept of “blue” in the Kazakh Language. *International Journal of Society, Culture & Language*, 11(2), 85-93. <https://doi.org/10.22034/ijscsl.2023.556420.2676>
- Potanin, G. N. (1972). *Kazakh folklore in the collection of G.N. Potanin*. Nauka.
- Radlov, V. V. (1899-1911). *Experience of the dictionary of Turkic dialects*. Academy of Sciences.
- Sevortyan, E. V., Levitskaya, L. S., Dybo, A. V., Rassadin, V. I., Sevortyan, E. V., Levitskaya, L. S., Dybo, A. V., & Rassadin, V. I. (Eds.). (1974). *EDTL – Etymological dictionary of Turkic languages*. Nauka.
- Sklyarenko, A. M., & Sklyarenko, O. M. (2018). *Typological onomastics*. Astroprint.
- Superanskaya, A. V., & Suslov A. V. (2008). *About Russian names*. Azbuka.
- Tatar language corpus. (2023). *Tugan tel national corpus of the Tatar language*. <https://tugantel.tatar/search/word/%D0%BF%D3%99%D0%B9%D0%B3%D0%B0%D0%BC%D0%B1%D3%99%D1%80/?groupBy=morph>
- Telia, V. N. (Ed.). (1996). *Russian phraseology: Semantic, pragmatic and linguocultural aspects*. Shkola.

- Tse, G. Y. W. (2003). Validating the logistic model of article usage preceding multi-word organization names with the aid of computer corpora. *Literary and Linguistic Computing*, 18(3), 287–313. <https://doi.org/10.1093/lc/18.3.287>
- Uali, N. (2014). Linguocognitive bases of disclosure of historical-cultural codes in onomastic names. *Onomastic Bulletin*, 2(28), 6-14.
- Vinogradova, L. N. (2000). *Folk demonology and mytho-ritual tradition of the Slavs*. Indrik.