opción

Revista de Antropología, Ciencias de la Comunicación y de la Información, Filosofía, Lingüística y Semiótica, Problemas del Desarrollo, la Ciencia y la Tecnología

Año 36, 2020, Especial N° 27

Revista de Ciencias Humanas y Sociales
ISSN 1012-1587/ ISSNe: 2477-9385
Depósito Legal pp 198402ZU45

Universidad del Zulia
Facultad Experimental de Ciencias
Departamento de Ciencias Humanas
Maracaibo - Venezuela

Opción, Año 36, Especial No.27 (2020): 1864-1900
ISSN 1012-1587/ISSNe: 2477-9385

Multilingual Thesaurus of Industry-Specific Terms as Major Aids for Translators

Ainur T. Bayekeyeva[1]
[1]Department of Translation Theory and Practice
L.N.Gumilyov Eurasian National University
Address: 010008, 11 Kazhymukan Street, Nur-Sultan, and Republic of Kazakhstan
a_baekeeva@mail.ru

Saule Zh. Tazhibayeva[2]
[2]Doctor of Philology, Professor
Department of Translation Theory and Practice
L.N.Gumilyov Eurasian National University
Address: 010008, 11 Kazhymukan Street, Nur-Sultan, and Republic of Kazakhstan

Aigul A. Shaheen[3]
[3]Candidate of Philology, Associate Professor,
Department of Theoretical and Applied Linguistics
L.N.Gumilyov Eurasian National University
Address: 010008, 11 Kazhymukan Street, Nur-Sultan, and Republic of Kazakhstan

Zhainagul S. Beisenova[4]
[4]Doctor of Philology, Professor
Department of the Russian Philology
L.N.Gumilyov Eurasian National University
Address: 010008, 11 Kazhymukan Street, Nur-Sultan, and Republic of Kazakhstan

Gulnar B. Mamayeva[5]
[5]Candidate of Philology, Associate Professor,
Department of Kazakh Linguistics
L.N.Gumilyov Eurasian National University
Address: 010008, 11 Kazhymukan Street, Nur-Sultan, and Republic of Kazakhstan

Recibido: 20-12-2019 •Aceptado: 20-02-2020

Abstract

This paper discusses the theoretical and methodological issues of compiling multilingual thesaurus for conducting industry-specific translation in the mining industry as one of the key genres of translation studies that bear difficulties in translating the technical documents. In market realism of Kazakhstan, industry-specific translations are one of the most popular types of translation services. It is reasonably considered that only a translator with technical background with the knowledge of the subject field is able to translate a technical text well, i.e. an industry-specific technical translator who understands the essence of the matter, the so-called 'engineering competent translation'.

Keywords: Thesaurus, Multilingual thesaurus, Term, industry-specific terminology, Mining and metal terms.

Tesauro multilingüe de términos específicos de la industria como ayudas principales para traductores

Resumen

Este documento discute los problemas teóricos y metodológicos de la compilación de tesauros multilingües para realizar traducciones específicas de la industria en la industria minera como uno de los géneros clave de los estudios de traducción que tienen dificultades para traducir los documentos técnicos. En el realismo de mercado de Kazajstán, las traducciones específicas de la industria son uno de los tipos de servicios de traducción más populares. Se considera razonablemente que solo un traductor con experiencia técnica con el conocimiento del campo temático puede traducir bien un texto técnico, es decir, un traductor técnico específico de la industria que comprende la esencia del asunto, la llamada 'traducción competente en ingeniería'.

Palabras clave: Tesauro, Tesauro multilingüe, Término, terminología específica de la industria, Minería y términos de metal.

1. INTRODUCTION

Industry-specific translation is a part of not only extralinguistics and intralinguistics, but also serves as the interdisciplinary science between the theory and practice of translation and industry-specific terminology. The mining and metal terms of the Kazakh language were first unified in 1950s by the initiative of the famous scientist Kanysh Satpayev. However, within the framework of the current global 'fourth industrial revolution', the mining potential of Kazakhstan is growing internationally, with the mining term stock being replenished with various new technology names and devices.

Most of the professional guides, dictionaries, and glossaries as for translating in the mining industry in Kazakhstan are systematized in alphabetical order and have several variants for each term. That is why there is a need to create a systemized online multilingual Kazakh-Russian-English thesaurus of industry-specific terms in order to freely use the industry-specific terms for providing adequate translation. The purpose of the research is to unify and systematize mining and metal terms in the Kazakh language according to the requirements of modern online thesaurus; to describe and provide Russian and English translations; to classify their differentiation into essential subjects and species.

In the research, the following methods were used: qualitative method, quantitative method, descriptive method, method of theoretical analysis of scientific data, method of classification and systematization. The research considers more than 10.000 mining terms in the Kazakh language, provided by the Terminological Committee of the Kazakh language. In addition, search tools for adequate translation have been demonstrated and proven with specific examples to provide industry-specific terms translation. Industry-specific texts are usually highly standardized, as they convey the information with the specific terminology, particularly mining and metal terms. Semantics and vocabulary should be equally accurate in both languages: SL and TL, i.e. the source language and the target language.

While translating the scientific and technical style of documents, i.e. industry-specific texts in Kazakh, Russian an English languages, the written documentation with certain additional requirements and features are demanded, namely, terminology and thesaurus. Terminology, as the independent branch of linguistics, concerned with the study and compilation of the terms of specific fields is not a new field for study, but during the recent decades, terminology has been systematically developed, with full consideration of its principles, bases and methodology. Terminology first began to take shape in the 1930s and has only recently moved from amateurism to a truly scientific approach (Cabrė, 1998).

Thesaurus is one of the major tools in the translation theory and practice, particularly in compiling various dictionaries, glossaries for

specific fields and purposes that requires the participation of the wide range of specialists in linguists, information technology, branch of science and technology, etc.

The term thesaurus has various definitions provided by different linguistic resources and dictionaries. Thesaurus (‘treasury, storehouse’), has the meaning of ‘treasure, treasury, storehouse’ (Latin) and ‘treasure, treasury, storehouse, chest’ (Greek). The notion is used since III-IV century in Amarakosha Sanskrit thesaurus (Online thesaurus). Thesaurus with the meaning of ‘collection of words arranged according to sense’ is first attested in Roget’s thesaurus, Thesaurus of English Words and Phrases, 1852. (Online thesaurus)

Thesaurus is a book of words or of information about a particular field or set of concepts; and Thesaurus is a list of subject headings or descriptors usually with a cross-reference system for use in the organization of a collection of documents for reference and retrieval. (Merriam-Webster’s dictionary; Collins dictionary)

Online thesaurus (thesauruses, thesauri) is a dictionary of words with the same or nearly the same meanings, or synonyms, and their opposites, or antonyms, such as thesaurus.com, available on the Internet or the World Wide Web, accessed through a web browser, and used by entering a query term into a search box on the site. An online thesaurus provides immediate electronic access to lists of alternate terms for the queried word, covering its various shades of meaning. (Online Dictionary of English)

Thesaurus in the modern applied sciences are considered to be as WordNet, EuroWordNet or Wiktionary linguistic resources,

describing the relationship between lexical meanings of natural language as a hierarchical system of synonymic groups, i.e. synsets (Tussupov et al, 2016; Tazhibayeva et al, 2016).

2. METHODOLOGY

Currently in the era of information technologies used in various fields of human life, particularly for translators in using machine translation, it is necessary to develop the industry-based thesaurus or multilingual thesaurus for specific field of study. Thin Digital Library e following study of compiling multilingual thesaurus of mining and metal terms has been conducted within the joint project between L.N. Gumilyov Eurasian National University (Kazakhstan) and Institute of Computational Technologies of Novosibirsk State University (Russian Federation).

Translating for mining industry starts from developing special terminological base and in order to solve the systematization of the mining terms, we decided to compile the industry-based thesaurus of mining and metals in Digital Library. The major purpose of compiling thesaurus is directly related to the information support of scientific research and translation activities, as well as with the formation and development of its own electronic resources for the translation industry and industry-specific terms, with the interoperable management of these resources publications and bibliographic resources. For ensuring information in industry-specific terms, in our case, the mining and

metal terms of the Kazakh, Russian and English languages, it was decided to use the universal Integrated Distributed Information System (IRIS) to compile the thesaurus developed by Institute of Computer Technology and Science of Siberian Branch of the Russian Academy of Sciences in 1998. (Fedotov et al, 2016) To date, the most effective way to solve lexicographic problems in the implementation of industry-based translation activities is to organize information about them into information systems exposed in the form of electronic libraries. In the framework of the thesaurus development, digital libraries are considered as a separate specific technology for working with digital information, forming a new class of information systems (IS) designed to manage information resources (Shokin et al, 2015; Lukashevich 2011).

The term digital library (DL) in our case is understood as a system for managing structured cataloged collections of dissimilar digital resources. DL not only provides a comprehensive search and navigation through catalogs (as opposed to print publications, microfilms and other media), but also provides the user with a directly found resource (publication, document, photo, fact description, etc.), as well as additional information about it. For example, geo-referencing, information about facts, bibliography, list of geological organizations, etc. The developed model of the information system was implemented in the form of the Digital Library Management System (IRIS) and has been used by Institute of Computer Technology and Science of Siberian Branch of the Russian Academy of Sciences since 2004. (See Fig. 1)

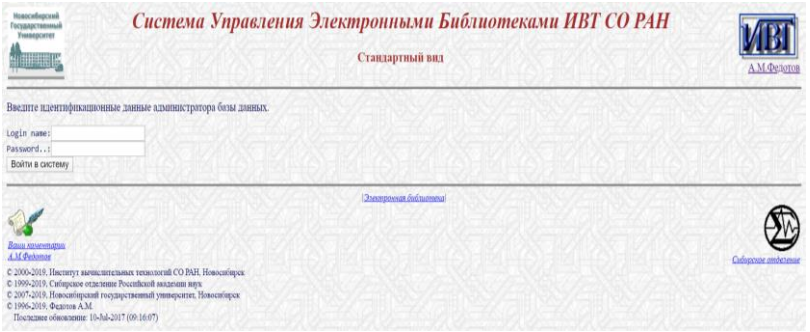

Figure 1: Homepage of Digital Library (Multilingual Thesaurus for Mining and Metal Terms)

The standard approach to the systematization of information is the classification of documents using taxonomies. Taxonomy is a subject classification that groups terms into a managed thesaurus and organizes the dictionaries into hierarchical structures. To describe a subject area, a certain set of key terms is usually used, each of which means or describes a concept from a given subject area. (Salton 1975, 1979) The basis of the classification is the allocation of concepts (key terms), the establishment of paradigmatic relations (for example, the type of parent - child) between them and the comparison of the analyzed document to the selected concepts. The problem in providing scientific and educational information systems for mining industry is that the technologies for classifying and systematizing information developed by libraries over the past hundred years do not work because of the thematic proximity of the classified documents.

According to the definition of an International Organization for Standardization (ISO), a thesaurus is a dictionary, managed indexing

language, formally organized in order to establish explicit a priori relationships between concepts. (ISO 1985; ISO 1986) This definition establishes lexical units and semantic relationships between these units as elements that make up the thesaurus. Thesaurus relations (genus - species, part - whole, etc.) are imposed on the taxonomy structure, i.e. identifies the main taxonomy of the subject area. Historically, the thesaurus was created for manual indexing of documents and when creating them, issues related to automatic indexing were not taken into account.

So, as we mentioned above, the thesaurus in the modern applied sciences are considered to be as WordNet, EuroWordNet or Wiktionary linguistic resources, describing the relationship between lexical meanings of natural language as a hierarchical system of synonymic groups, i.e. synsets. The difficulty of compiling a thesaurus that corresponds to the whole thematic variety of indexed information is the main reason for its unpopularity in modern information systems. However, the effectiveness of information retrieval systems to support scientific and educational activities directly depends on the use of specialized thesaurus. One of the first ever thesauruses (ideographic dictionaries) most widely known today was compiled by the British lexicographer Peter Marc Roger and published in 1852. The original name of the Roger Thesaurus is Thesaurus of English Words and Phrases.

In the field of machine translation for the first time the thesaurus was used by M. Masterman. In 1961, a paper appeared in which it defined the basic vocabulary for 15,000 concepts. With the help of

thesauruses, a correspondence was established between the language of user requests and documents in the information system. (Masterman, 1961) Schreider Yu.A. Proposed to consider the thesaurus as a system of knowledge reflected in the language, then the thesaurus becomes interesting in itself, and not only as an auxiliary tool. (Schreider, 1963)

Among universal automated thesauruses, WORDNET intelligent electronic thesaurus for English, developed at Princeton University, and the analogue RussNet thesaurus for Russian, developed by Mathematic Linguistics Department of St. Petersburg State University, and multilingual UNESCO Thesaurus. Information retrieval thesaurus is a normative and controlled dictionary of key terms in natural language with explicitly indicated semantic relations between terms, intended to describe the content of documents and search queries. (ISO 2011; ISO 2013) The thesaurus is intended to describe a specific subject area, each term of which denotes or describes a concept from a given subject area.

The terms of mining and metallurgy constituting the thesaurus are divided into descriptors and ascriptors. Descriptors are uniquely consistent with the concepts of the subject area. Relations between descriptors are usually divided into two types: hierarchical and associative. Hierarchical relationships are usually considered as asymmetric and transitive.

In accordance with GOST 7.25-2001 (GOST, 2001), hierarchical relations have the properties of transitivity and ant symmetry, which can be used for excessive indexing in the interests of

increasing the efficiency of information retrieval. The hierarchical relations used in information retrieval thesauri can be differentiated into separate types. The main hierarchical relation used in information retrieval thesauri is the generic relation (parent – child). A generic relationship is established between two descriptors, if the scope of the notion of a descriptor is included in the scope of the notion of a higher descriptor. Also, as part of a hierarchical relationship in information retrieval thesauri, a 'part-whole' relationship can be established. The relation of association is non-hierarchical and associative. The main purpose of establishing associative relationships between information retrieval thesaurus descriptors is an indication of additional descriptors useful in indexing or searching. (ISO 2013)

The information retrieval thesaurus (IR thesaurus) model described in national and international standards is intended for its use in the process of manual, expert analysis of documents. The main purpose of developing traditional information retrieval thesauri is to use their units (descriptors) to describe the main topics of documents in the process of manual indexing. Therefore, it is important that the set of descriptors of the information retrieval thesaurus makes it possible to describe the subject matter of the domain documents. At the same time, the very process of indexing according to such a thesaurus is based on linguistic, grammatical knowledge, as well as knowledge of the subject area, which is available to professional text indexers. The indexer must first read the text, understand it, and then set out the content of the text, using the descriptors specified in the information retrieval thesaurus. The indexer must have a good

understanding of all the terminology used in the text to describe the main topic of the text, he will need a much smaller number of terms.

IR thesaurus for mining and metal terms is constructed to describe the vocabulary of descriptive IR language, the lexical units of which are descriptors. The following are the notions basically used in constructing thesaurus for information retrieval purposes:

Lexical unit of the IR language is the designation of an individual concept, adopted in the IR language and indivisible in this function.

Notes are the lexical units that may be accepted in the natural language, set phrases or collocations, generally accepted abbreviations and acronyms, symbols, dates, lexically significant components of complex words, as well as equivalent code or symbolic notations for artificial language, for example codes of classes of the classification system (Laftimi, 2012). Descriptor is a keyword selected from a group of conditionally equivalent keywords and represents this group when indexing and searching for information. The descriptor is also described as a semantic dominant, or a basic concept with a word related to it, like a header word in an explanatory dictionary. Terms can serve as descriptors that denote certain concepts of a certain subject area and satisfy the principles of common usage, prevalence, brevity and terminological accuracy.

Keyword is a separate word or phrase of natural language, extracted from the text of the information document and reflecting its main content when indexing. The group of conditionally equivalent keywords unites not only those words and phrases that are recognized

as synonyms in the natural language, but also those that can be considered conditionally equivalent in terms of information retrieval, i.e. within the framework of the given IR language. In our case 3 IR languages: Kazakh, Russian and English.

Paradigmatic relations (basic, analytical and associative relations) express the constant semantic relations between lexical units of IR language, not depending on the text. The following relations are called the semantic relations between lexical units of IR languages: 'class-type', 'whole-part', etc. These lexical units are stable for each subject area and can be fixed in the dictionary. E.g., trammeling-грохочение-тасуату, crushing-дробление-ұсату, grinding-измельчение-ұнтау, flotation-флотация-флотация, filtration-фильтрация-сүзгілеу, magnetic separation-магнитная сепарация-магнитті сепарациялау are related to category 'Enrichment-Обогащение-Кен байыту', i.e. the separate terms as flotation or filtration in 'class-type' relation between the descriptors of enrichment may be independent in the context. This is called paradigmatic relation in denoting the descriptors.

In IR thesaurus the following 3 paradigmatic relations are usually fixed and applied: class-type, equivalence/synonymic and associative relations.

(i) The 'class-type' relation links two descriptors if the scope of the concept corresponding to one of the descriptors includes the scope of the concept of another descriptor, e.g. Crushing Department – crushing workshop (Дробильный цех – дробильная; Ұсату цехы – ұсату бөлімшесі), Crushing Unit – crushing machine

(Дробильная техника – дробилка; Ұсату техникасы – ұсату машинасы), Crushing Type – medium crushing, fine crushing (Тип дробления – среднее дробление, мелкое дробление; Ұсату типі – орташа ұсату, ұсақ ұсату).

(ii) The synonymic relations mean that the search for one of the conditional or true synonyms will allow to find in the database of the Automated Information System (hereinafter referred to as AIS) those documents to which the others are assigned as keywords, e.g. enrichment = beneficiation. Both terms have the same meaning as обогащение-кен байыту.

(iii) Associative relations are established between keywords, belonging to the same or different semantic categories and arbitrary levels of the hierarchy. They are analogous to associative relations in the mind of a person/translator, when the emerging idea of one object causes ideas about others, e.g. when searching for a descriptor for metal production-производство металла-метал өндіру, the user/translator of AIS can be offered to conduct additional searches by descriptors: metallurgy-металлургия-металлургия; industry-промышленность-өндіріс; raw materials-сырье-шикізат; deposits-месторождения-кен орны; metals-металлы-металдар; enrichment/beneficiation-обогащение-кен байыту; processing methods-способы обработки-өңдеу тәсілдері; equipment-оборудование-құрылғы; transport-транспорт-көлік and etc.

As a rule, associative relations are not distinguished in IR thesaurus by their semantics. However, it makes sense to mention the

following types of associations: 'whole-part', 'cause-effect', 'proximity in space or time', 'antonymy', 'object-subject area', etc.

Ontology in terminology is a developed formalized means of describing the terms of the subject area that can be used in modern intelligent information systems. Ontology consists of a set of concepts and a set of statements about the following concepts: the classification of concepts, the relations between concepts, the hierarchy of concepts on the relations 'general-part' and 'part-whole'. Synset, or synonym set, represents a line of a WordNet pos data file. A Synset represents a concept, and contains a set of Words, each of which has a sense that names that concept (and each of which is therefore synonymous with the other words in the Synset) (Information and documentation, 2013). Each synset is supplemented by a definition and examples of the use of words in the context for convenience. A word or phrase can appear in more than one synset and have more than one category of part of speech. Each synset contains a list of synonyms or synonymic phrases and pointers describing the relations between other synsets. Synsets in WordNet are interconnected by various semantic relations: (i) Hyperonym: metal product → metal production; metal product → metallurgy; (ii) Hyponym: metal production → metal product; metallurgy → metal product; (iii) Has-member: metallurgy → metallurgist; (IV) Member-of: metallurgist → staff; (v) Meronym: has-part: enrichment → flotation; enrichment → filtration; (vi) Antonym: magnetic separation → non-magnetic separation.

In accordance with the definition of standards, IR thesaurus is a normative dictionary that accurately indicates the relationship between

terms and is intended to describe the content of documents and search queries and IR thesaurus is designed for the following objectives: to provide the translation of documents and user requests to the same dictionary applied for indexing and searching. Thus, the differences in the lexical composition of the document and the user's request are reduced to the same units of the thesaurus; to ensure the consistent use of index units; to provide relations between terms, i.e. relations between thesaurus units allow the users/translators to find the optimal term for describing a document or request; to use as a search tool when searching for documents.

So, when forming IR thesaurus, the first task is to select terms for including in the thesaurus. There are several possible sources of terms for forming IR thesaurus. Principally, existing thesauri should be studied in close subject areas. Thesauri can contain a considerable number of useful terms for the new thesaurus. Candidate terms for including in the thesaurus may be offered by subject matter experts. In addition, terms for thesaurus may be derived from domain-specific texts using automated methods or manual processing of documents. When manually processing documents, indexers index the incoming documents with the most relevant keywords, which are then reduced to a single list that may serve as the basis for the thesaurus. (Arkhangelskaya et al, 2001)

After the list of candidate terms is obtained, too-frequency terms are excluded from the list, since it is assumed that they are not informative enough to differentiate the separate documents. Relatively low-frequency terms can be removed from the list or presented as

ascriptors of more general or more frequency concepts. Gerd proposes to exclude some specific terms, since the thesaurus, in which many levels of hierarchy, is difficult to use: the subjectivity of indexing increases, since indexers can use descriptors of various levels for indexing documents. (ANSI/NISO, 2005) In case if there are several similar terms in the list, then the most frequency term is allocated, the remaining terms can be excluded and included in ascriptors.

The Legislative Indexing Vocabulary (LIV) of the US Congressional Research Service (US CRS, LIV, 1998) describes the rules for including terms in the thesaurus as follows: Shirley Loo, Legislative Indexing Vocabulary: The CRS Thesaurus (22nd ed.) i (1998)

(iv) Terms of Thesaurus should represent the concepts that are actually mentioned in the source document or literature and should be selected from considerations of the effectiveness of their use in the search for documents;

(v) An important factor in including the term is the frequency of its mention in the texts that must be checked periodically;

(vi) Including new terms in the thesaurus should take into account already included thesaurus terms. Candidate terms should be checked for consistency of their generality and specific use with other terms of the thesaurus. It should also be checked whether the candidate term represents a separate concept that does not have correspondences among the existing terms of the thesaurus. It is necessary to avoid including terms whose values overlap with the values of already

existing thesaurus terms so that it will be difficult for indexers, users and translators to differentiate the terms.

3. RESULTS and DISCUSSION

The development of terminology in historical language study, inter-language relations have always been in the focus of researchers' attention. At different times, the problems general terminology, field terminology and other issues of terms are in the sphere of interest of such researchers such as: A. Baitursynuly, M. Teresa Cabrė, A. Kaidar, Sh. Kurmanbayuly, Zh.S. Beisenova, and many others.

Translating industry-specific texts is achieved by the methods differed from the translation of literary works. A literary translator may be perceived as a portrait painter who aspires to portrait, rather than photographic similarity. Whereas a technical translator of scientific and technical literature is a photographer whose quality of work depends only on the accuracy of the transferred image. It can be said that the translation of fiction refers to the field of art, and technical translation refers to the field of science. According to the recent studies on term building, most scientists note that genuinely international terms that have the same meaning, especially in the field of industry, science and technology become more and more due to the international cooperation of professionals and scientists who develop coordinated international terminology at congresses and conferences.

As a result of interaction and mutual influence of national languages, the number of borrowings increases, and the modern terminological base of the Kazakh language is no exception. Moreover, there is a tolerant attitude of society towards borrowed words and terms, which is characterized by a close relationship with other nations in all areas of professional and social life. In recent years, besides the popular oil production, Kazakhstan has been strengthening and maintaining international relations in mining and metallurgy, and this may be called as a period of the new industrial wave, which entails the assimilation of new borrowed words and terms due to the new technologies used for the mining industry.

Language contacts, the rapid development and the presence of various forms of the Interlingua community should be denoted by the concept 'Internationalization of language'. Some scholars believe that this concept is able in a number of contiguous languages to emphasize the similarity of internationalism, both in terms of content and expression. (Isakova 2018: 40-46) Internationalization of language occurs when a large number of international units: words, morphemes, phraseological units, syntactic constructions in a language; and a large number of international words belong to scientific and technical terminology. (Gazda, 2002; Manerko et al, 2014) International terms exist in any human language, and the international terms enable to enrich the national language, make the TL flexible and in no way impair its originality since the lexical and grammatical structure inherent in a given language, its internal laws of language development and national identity are fully preserved.

In the Kazakh language, the borrowing of foreign terms occurs mainly through the English and Russian languages. However, some scholars have a critical attitude to borrowing and stand for a national language, not international terminology. At the same time, most scholars are not so categorical about borrowing, believing that the main requirement for borrowing terms is not to use the words of foreign languages if the familiar and motivated term with a similar meaning is already functioning in the native language.

Terminological lacunae is desirable to fill the existing internationalism. In the case of a different level of motivation of synonymous terms, preference should be given to the most motivated terms. (Taranova et al, 2016) This attitude to borrowing is more characteristic of translation theory of terminology in Kazakhstan. Regarding terms, we speak about full equivalence, as long as the term is mono semantic, context free and in the specialized field denote one concept. There are relations between concepts within one language, described in conceptual systems, and of course, between two or more languages, resulting from the comparison of the corresponding conceptual systems.

The reasons why Kazakh terminology is rich in international terms is that their use gives a rigorous scientific text solidity and greater scientific weight. This remark can rightly be attributed to any national terminology. However, we should not forget that the majority of scientists do not object to the use of international terms, noting that when borrowing, the measure should be kept, and the calculation itself should be motivated.

The set of terms of a particular field of industry, science and technology is called professional terminology or industry-specific terminology, and each field has its own terminology in accordance with the subject and method of its work. Terminology is the core of the scientific style. It can be said that the term embodies the main features of the scientific style and is extremely consistent with the tasks of scientific communication.

As other industry-specific terms, the mining and metal terms should provide a clear and accurate reference to real objects and phenomena; establish an unequivocal understanding of the transmitted information by professionals. Therefore, a special requirement is imposed on this type of words. While considering the terms, we have to bear in mind the following facts: First of all, the term has the exact meaning in each field of study, i.e. to have a strictly defined meaning, which can be revealed by a logical definition, establishing the place of the concept designated by the term in the system of concepts of a certain industry, science or technology. (Cabrė, 1998; Kaidarov, 1993; Kurmanbaiuly, 1997; Beisenova, 2011, 2014)

In English, there are unambiguous terms, such that express only one concept in any branch of industry, science or technology. Along with them, there are both homonymous and synonymous terms that completely coinciding with each other in external form, call different objects in different branches of industry, science and technology, or having one and the same meaning with various terms: enrichment and beneficiation has the same meaning of treating the ore in order to get various types of metals.

Gradually, the content of technical knowledge of the source language (SL) begins to penetrate the signs of the other language, enrich the vocabulary of the target language (TL). In any language, the word or the phrase is already inseparable from their meaning, and here the content of technical knowledge becomes an element of the language in a certain field, in our case the mining industry. The technical knowledge, which has found its expression in a word, in a term of SL, passes into a qualitatively new stage, joining the semantic structure of TL, becoming a component of the terminological system of the given language.

The processes of forming new terms in the modern English mining industry are associated with the emergence of new industrial, technical and scientific ideas and research fields, the development of new technology, the creation of new technological processes. For this reason, the new terms absent in the dictionaries of TL are usually carriers of the most important information for the recipient of the message, in order to extract which they are processed and studied. Much attention is paid to the systematic nature of the newly created terms. In many areas, special rules have been developed for the formation of terms for concepts or objects of a particular class. According to the classification of term building in the English language, the most intensive way of enriching the terminological composition of the word occurs via the following processes:

(i) Affixation, i.e. the formation of new single-word terms by adding prefixes and suffixes borrowed from Greek and Latin languages to the roots. In English, the addition of the prefix pre is used

to convey the idea of advancing in time or position: predesign is a draft design or Pre-FS (Pre-Feasibility Study) is a draft or primary Feasibility Study of any project. The prefix inter is derived from the Latin word inter, meaning between, among interchangeability. The prefix Trans means through, beyond transition. The suffixes have a broader meaning, indicating the category to which the term is being defined. For example, using the suffixes -er, -or, -ist, nouns are formed to indicate a specialist worker: steelworker, metallurgist, as well as machines, tools and instruments: cutter, jaw crusher, dozer, pressure filter, rock-breaker, cornish feeder. With the help of suffixation, new terms and words can be formed that fulfill the role of a noun, adverb or verb in a sentence.

The suffixes and prefixes used in the English term formation system are mainly borrowed from common word-forming means of the English language. For the formation of terms with the help of suffixes and prefixes, the choice of the number of derivational elements is those that are acceptable for the construction of terms. Specific to special terminology is the desire to assign certain terminological meanings to certain suffixes. This is especially characteristic of metallurgical, chemical and geological terminology, where the specialization of suffixes is most fully studied. In English, the following are formed: the names of basic organic compounds and halogens formed with the suffix -ine [en] (amine, oxygen); names of non-basic organic compounds formed with the suffix -ine [in], -yne (alcaline, propyne). English prefixes dis-, en- are widely used in the formation of mining terms: dislocation, enrichment.

(ii) Conversion of a word or term, with the help of which a new term is derived from an existing word or term relating to another part of speech, for example: to know+how=know how (know-how production technology). In English, in many cases, new terms are formed from words already existing in the language without any change in their spelling or pronunciation. The most common type of conversion is the formation of verbs from nouns, so from the noun generator the verb to generator (to work in the generator mode). In modern English, the formation of verbs from nouns is widespread and serves as one of the sources of the appearance of neologisms.

(iii) Direct borrowing of terms from Greek and Latin languages, for example, locus (geometric location of points); versus (depending on how the function of any argument).

(iv) Formation of complex terminological groups by adding various definitions to the original term: evaporation system. This term is highly specialized and unambiguous, as it is used only in metallurgy and enrichment and has no homonyms. It should be noted that this prevailing method of term formation by adding definitions to a term in the industrial, scientific and technical literature is explained by the tendency of modern science and technology to create complex systems and objects, which in turn consist of sets of constituent elements.

(v) Shortening and blending the words or phrases in the specific field. Recently, the tendency to form new terms by shortening existing words and phrases has intensified. The growth of abbreviations and acronyms is explained by the fact that complex

words and phrases are inconvenient in professional use and, naturally, there appears a desire to convey them briefly. Abbreviations and acronyms are both shortened versions of word combinations and expressions where both can often be represented as a series of letters. An abbreviation is any shortened or contracted form of a word or phrase: Mon. (Monday), Jan. (January), TechSpec (Technical Specification), etc. An acronym is a specific type of abbreviation formed from the first letters of a multi-word term, name, or phrase, with those letters pronounced together as one term: FS (Feasibility Study), JORC (Joint Ore Reserves Committee), MPP (Mining and Processing Plant), etc. (Bayekeyeva, 2019)

In many industrial areas, special rules for the formation of terms for concepts or objects of a certain class have been developed: the names of various types of rocks and minerals are created by analogy with the suffix -ite, indicating the origin of the materials: andesite, dacite, diorite, pegmatite, peridotite, quartzite, and etc.

Based on the foregoing, it can be concluded that the mining and metal terms have evolved over a long time and has now been formed into a strict terminological system. Within this system, certain forms of term building have been formed, the degree of productivity of which makes it possible to determine the dynamics of growth of terms and further ways of their development. In connection with the progress of industry, science and technology, where new concepts are constantly being opened, for the designation of which it is necessary to use more and more new terms in the English, Kazakh and Russian languages. Term building takes place both at the expense of the resources of the

English language itself and by borrowing from other languages. This process lets the language develop, makes the terminology international, which allows scientists and professionals to understand better each other.

The research is aimed at the selection of the optimal translation of the mining and metal terms in the specific industry. The problem of translator's 'false friends' can be applied to the translation of scientific and technical texts; it has its characteristic peculiarities. It is possible to note that the translator's 'false friends' in the scientific and technical texts of mining industry are characterized by certain specific properties. Some words related to the translator's 'false friends' category are considered to be international words.

The major difficulty of translating industry-specific technical texts is the accurate transferring of special terms without misunderstanding. This is impossible without the technical background of a technical translator and understanding the meaning of the translated text. Because the mining industry is a specific type of translation and has variations in term use with the oil and gas industry, though both of industries have similar geologic terms in exploration and production of natural resources.

For instance, while searching for the English equivalent for the Kazakh ұңғыма and Russian скважина, one can find several synonyms for this term as bore, borehole, whole and well. The translator will have a problem in choosing a certain correct form of the term between borehole and well. Let's present an analysis of the above-mentioned terms. According to the explanatory dictionary for

well drillers, 'Typically a borehole is drilled by machine and is relatively small in diameter. A well is usually sunk by hand and is relatively large in diameter'. However, this is not enough for a technical translator for the mining industry, and borehole should be chosen as the proper equivalent for the mining industry, because well is used only for oil and gas industry because of the liquid nature of oil, and borehole is used for mining industry because of the solid nature of the ore. So, every of the four English terms has its own particular meaning in a precise sphere. However, in Kazakh and Russian, it is used with respect to only one meaning ұңғыма and скважина which has no specific semantic varieties.

One more difficulty in translating for the mining industry is the widely used synonymic range of terms. For example, while translating the English text, a translator may meet the terms as beneficiation, concentration, dressing, enrichment, refinement, separation, treatment, and washing for the Kazakh кен байыту and Russian обогащение, and the translator will have a problem in choosing a certain correct form of the term between the ranges of synonyms. Here comes the technical background of the translator: concentration is used for concentrating the chemicals in laboratory; refinement is used only for gold and silver refining in mining and metallurgy (but oil can also be refined); separation is the exact process of separating the different metals in the ores; washing is the initial process of separating the ores; dressing is used for treating the ore in the special equipment; and treatment is the general name for processing the ores of any type. Only

enrichment and beneficiation may be used synonymously for the Kazakh кен байыту and Russian обогащение.

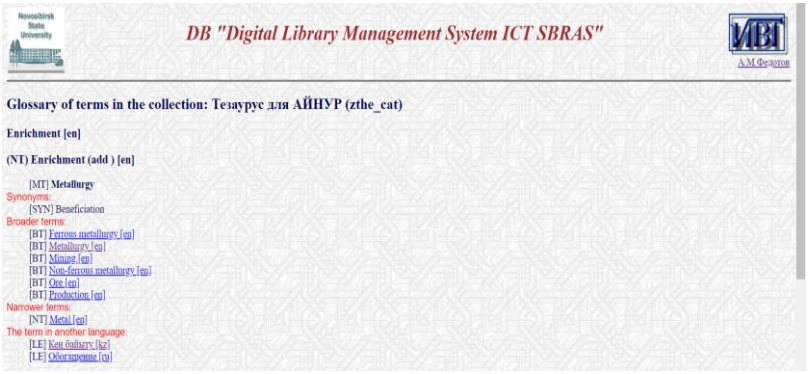

Figure 2: The term ‘enrichment’ in Multilingual Thesaurus for Mining and Metal Terms

The main advantage of the given using multilingual thesaurus for mining and metal terms is that the thesaurus will give the most adequate option of any term with maximum one perfect synonym. As shown in Figure 2 and Figure 3, the term enrichment will be given with the broader and narrower terms applied for mining and metal industry (See Fig. 2, Fig. 3). In its turn, the broader term will give the general description for the term relied.

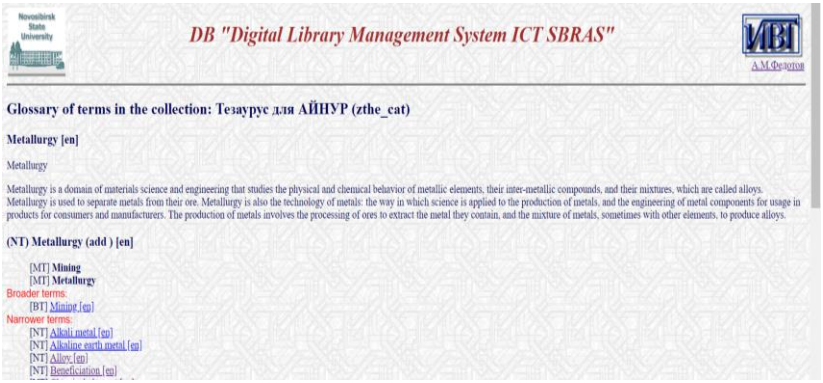

Figure 3: The broader term and its description in Digital Library (Multilingual Thesaurus for Mining and Metal Terms)

The first stage in translating mining terms involves studying the meaning of the source-language mining term to be translated. Then, after having compared the technical background studies, a term with the same content must be sought in the target-language mining system. Equivalence aims to give the lexis and terminology of two languages equal meaning and corresponding import and significance, and, as can be seen from some of the research presented above, it also strives to achieve the same terminological effect based on technical interpretation of the information in the source language.

Therefore, to achieve professionalism in translating technical data of mining industry, a technical translator for mining industry should master a required number of field-specific terminology, denoting the basic concepts of geology, mining and metallurgy for definite stages of workflow in exploration, production, beneficiation, and enrichment.

4. CONCLUSION

Mining industry in Kazakhstan has a crucial role in providing sustainable development and ensuring the effective operation of subsoil use in the following fields: Exploration, development, production, processing and marketing of solid minerals; Development and introduction of new high-tech and efficient technologies in mining industry; Reproduction of mineral resources potential of Kazakhstan; Development of commercial and non-commercial reserves deposits.

One of the problems of the translation business, like any other, is the lack of highly qualified professionals. It is the most acute for companies striving to provide high-quality services, because it requires professional translators and editors. Moreover, if the company specializes in technical translations of complex technical documentation, the situation becomes critical. Now there are quite a lot of specialists on the labor market with Translation, linguistic or philological background, but experience suggests that high-quality translation of highly specialized technical documents is beyond their power, and the need in highly qualified professional in the field of technical translation remains demanded.

Translation of technical data is a difficult and laborious work that requires a translator high qualification and knowledge of a certain area of translation: a technical translator should master the language of technical literature both in SL and TL. The prerequisites for a Technical Translator are the sufficient linguistic and general technical knowledge received during the study at higher education institutions.

In addition, a Translator needs to be aware of high language acquisition, academic writing skills and have a good sense of both SL and TL, especially the central requirements lie on getting in-depth knowledge in the required subject area and skills of technical thinking, and one of the important type is abbreviations and acronyms as part of the technical terms. In general, the adequacy of translating the abbreviations and acronyms is a full transmission of the semantic contents of the source text and a complete functional and stylistic correspondence in the target text. Therefore, analyzing the abbreviations and acronyms frequently used in mining and metallurgy in Kazakhstan, we can steadily say that the shortened form of the specific terminology can be called as one of the varieties of field-specific terms.

Based on the analysis of standards and various approaches to the implementation of thesaurus, a decision was made to use the Zthes data scheme for creating a thesaurus for mining industry. Currently, within the framework of the thesaurus of mining and metal terms, the major rubrics are formulated and about 10,000 Kazakh terms for mining and metallurgy have being analized and constantly updated. The use of the thesaurus in digital libraries is most effective with the continuous modernization of the thesaurus, its integration into the database and the corresponding level of thematic specialization. For the time being, the main use of the thesaurus is navigation through library resources and classification or categorization.

As a major translator's tool, thesaurus (from Greek θησαυρός - treasure) is a collection of information (corpus, arch) encompassing

with maximum fullness of the concept, definitions and terms of a special field of knowledge or industry, with examples of their use in texts. In modern linguistics thesaurus is a special type of vocabulary of general or special vocabulary, in which semantic relations (synonyms, antonyms, paronyms, hyponyms, hyperonyms, etc.) between lexical units are indicated.

Thesaurus of mining and metal terms is composed by at least three elements: (i) A list of words (or terms); (ii) The relations between the words (or terms), indicated by their hierarchical relative position (e.g. metal – broader term; iron concentrate – narrower term, iron content and Fe-content – synonym, etc.); (iii) A set of rules on how to use the thesaurus.

Thus, in order to form the up-to-date and good Information Retrieval Thesaurus for Mining industry, it is necessary to provide the following basic criteria: (i) A set of descriptors of the thesaurus should be sufficient to describe an arbitrary domain document, particularly for 'Mining and Metallurgy' industry; (ii) The number of descriptors should not be too large.

Studying the topical issues concerning thesaurus, we can steadily say that online thesaurus for industry-specific terms is one of the effective tools for describing individual subject areas and are designed to promote understanding in communication and interaction of professionals and translators for mining industry. Unlike the explanatory dictionary, the thesaurus makes it possible to reveal the meaning not only with the help of definition, but also through the correlation of the word with other concepts and their groups, so that it

can be used to fill the knowledge bases of artificial intelligence systems.

REFERENCES

AITBAIULY O. 2013. **Qazaq til biliminiń termınologıasy məseleleri**. Almaty: «Abzal-Aı» baspasy. – 400 p.

ARKHANGELSKAYA V.A. 2001. **Informatsionno-poiskovyi tezaurus po ekonomike i demografii / Ed**. Mdivani R.R. Moscow, 492 p.

BAISENOVA ZH. 2014. **Intercultural Approach of Professional Lexicon in Veterinary Medicine**. Aitmatov Academy, London – 172 p. ISBN 978-0-9926186-5-0.

BEISENOVA ZH.S. 2011. **Otraslevaya terminologiya: sistemnost, tipologiya, funktsionirovaniye**. Monograph, 2nd edition. Astana - 228 p. ISBN 938-601-7163-02-0.

BAYEKEYEVA A.T. 2019. **Technical Translation: Translating Abbreviations and Acronyms for Mining Industry**. Bulletin of Pavloar State University named after S. Toraigyrov, №1, Pavlodar.

CABRÈ M. TERESA. 1998. **Terminology: Theory, Methods and Applications. John Benjamins Publishing Company**. Amsterdam/Philadephia.

GAZDA J. 2002. **Internatsionalizatsiya kak integriruıyushii element yazykov sredneevropeiskogo prostranstva // Crossroads of Cultures: Central Europe**. Literaria Humanitas. Digital Library of the Faculty of Arts, Masaryk University. №011-2002-01_07.

ISAKOVA S. 2018. **Kirme termınderdi qabyldaýdyń sheteldik tájirıbesi**. «Termınjasam qaǵıdattary jáne ulttyq termınologıa máseleleri» respýblıkalyq ǵylymı – teorıalyq konferentsıa materıaldary, Astana. 40-46 pp. URL: https://termincom.kz/articles/?id=74

KAIDAROV A. 1993. **Qazaq termınologıasyna jańasha kózqaras**. Almaty - 42 p.

KURMANBAIULY SH. 2014. **Qazaq termınologıasy: zertteýler, oqýlyq, sózdik, bıblıografıa**. Almaty: Sardar. - 928+16 p. ISBN 978-601-7174-41-5.

LAFTIMI I. 2012. **Informatsıonno-poıskovye tezaurusy: osnovnye ponyatiya, naznacheniye i metodika razrabotki**. Otraslevoi rybolovnyi tezaurus // Molodoi uchenyi. №7, 164-166 pp.

LUKASHEVICH N.V. 2011. **Tezaurusy v zadachah informatsionnogo poiska. Moscow**: Moscow State University.

MASTERMAN M. 1961. **Semantic message detection for machine translation, using an Interlingua**. Proceedings of International Conference on Machine Translation. - 438-475 pp.

MANERKO L.A., KLAUS-DIETER BAUMANN, HARTWIG KALVERKÄMPER. 2014. **Terminology Science in Russia today: From the Past to the Future**. Frank&Timme. ISBN 9783732900510.

SALTON G. 1975. **Dynamic information and library processing**. – N.J.: Prentice Hall / Russian translation 1979: Динамические библиотечно-информационные системы. Moscow, Mir.

SHREIDER YU.A. 1963. **Quantitative characteristics of semantic information.** Research Institute. - Series 2, № 10, 35-39 pp.

SHOKIN YU.I, FEDOTOV A.M, ZHIZHIMOV O.L, FEDOTOVA O.A. 2015. **Evolyutsiya informatsionnyh sistem: ot Web-saitov do sistem upravleniya informatsionnymi resursami**. Bulletin of Novosibirsk State University. Information Technologies series. V. 13. №1.

TARANOVA YE.N, YERMAKOVA L.R, TARANOV A.O. 2016. **Internatsionalnye terminy kak obekt lingvisticheskogo issledovaniya (na primere terminov landshaftovedeniya v russkom i nemetskom yazykah)**. Nauchnye vedomostı BelGU. №21 (242). Volume 31. URL: https://cyberleninka.ru/article/v/internatsionalnye-terminy-kak-obekt-lingvisticheskogo-issledovaniya-na-primere-terminov-landshaftovedeniya-v-russkom-i-nemetskom

TAZHIBAYEVA S, YESKELDIYEVA B. 201Multilingualism in modern Kazakhstan: New challenges // Asian Social ScienceVolume 11, Issue 6, 25 February 2015, Pages 56-64.

TUSSUPOV J., SAMBETBAYEVA M.A., FEDOTOV A.M., SAGNAYEVA S.K., BAPANOV A.A., NURGULZHANOVA, A.N.YERIMBETOVA A.S. 2016. **Using the thesaurus to develop it inquiry systems**. Journal of Theoretical and Applied Information Technology.-Volume 86, Issue 1, 10 April 2016, Pages 44-61

TUSSUPOV J., SAMBETBAYEVA, M.A., FEDOTOV, A.M. FEDOTOVA, O.A.SAGNAYEVA S.K. BAPANOV, A.A., TAZHIBAEVA, S.Z. 2016. Cl**assification model and morphological analysis in multilingual scientific and educational information systems.** Journal of Theoretical and Applied Information Technology. Vol. 86, Issue 1, 10 April 2016. - pp. 96-111

Standards and Dictionaries

ANSI/NISO Z39.19: 2005 Guidelines for the construction, format and management of monolingual controlled vocabularies. NISO Press: Bethesda, MD, 2005.

ANSI/NISO Z39.50-2003. Information Retrieval (Z39.50): Application Service Definition and Protocol Specification. NISO Press, Bethesda, Maryland, U.S.A. November 2002.

ANSI/NISO Z39.19-2005, Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies. URL: http://www.niso.org/standards/resources/Z39-19-2005.pdf

IFLA. Guidelines for multilingual thesauri. (IFLA professional (IFLA professional reports:115). IFLA, 2009. URL: http://www.ifap.ru/library/book411.pdf.

ISO 5964:1985. Documentation - Guidelines for the establishment and development of multilingual thesauri. Geneva: International Organization for Standardization.

ISO 2788:1986. Documentation - Guidelines for the establishment and development of monolingual thesauri. 2nd ed. Geneva: International Organization for Standardization.

ISO 25964-1:2011. Information and documentation - Thesauri and interoperability with other vocabularies - Part 1: Thesauri for information retrieval.

ISO 25964-2:2013. Information and documentation - Thesauri and interoperability with other vocabularies - Part 2: Interoperability with other vocabularies.

ISO/IEC 13250:2000. Information technology - SGML Applications - Topic Maps.

ISO/IEC 13250:2003. Information technology - SGML Applications - Topic Maps.

ГОСТ 7.25-2001. Information retrieval thesaurus, monolingual. General rules, structure and composition / Interstate council on standardization, metrology and certification. Moscow, Standardinform.

COLLINS DICTIONARY. URL: https://www.collinsdictionary.com/

EPMG - Levels of Technical Translation Quality. URL: http://www.epmg.kz/ru/epmg/quality.html

EXPLANATORY DICTIONARY FOR WELL DRILLERS. URL: https://www.welldrillers.org/faq/about-boreholes-and-wells/

MERRIAM-WEBSTER'S DICTIONARY. URL: https://www.merriam-webster.com/

OFFICIAL WEB-SITE OF WORDNET. URL: https://wordnet.princeton.edu/

OFFICIAL WEB-SITE OF UNL MIR BY UNDL FOUNDATION. URL: http://www.unlweb.net/unlweb/index.php?option=com_content&view=article&id=57:unl-mir&catid=1:latest-news&Itemid=60

ONLINE THESAURUS. URL: http://www.thesaurus.com

ONLINE DICTIONARY OF ENGLISH. URL: http://www.dictionary.com/

TERMINOLOGICAL BASE OF THE REPUBLIC OF KAZAKHSTAN. URL: https://www.termincom.kz/termins

UNESCO THESAURUS. URL: http://vocabularies.unesco.org/browser/thesaurus/en/

UNIVERSIDAD DEL ZULIA

opción
Revista de Ciencias Humanas y Sociales

Año 36, N° 27, (2020)

Esta revista fue editada en formato digital por el personal de la Oficina de Publicaciones Científicas de la Facultad Experimental de Ciencias, Universidad del Zulia.
Maracaibo - Venezuela

www.luz.edu.ve

www.serbi.luz.edu.ve

produccioncientifica.luz.edu.ve