

REPUBLIC OF AZERBAIJAN

On the rights of the manuscript

ABSTRACT

of the dissertation for the degree of Doctor of Philosophy

**SEMANTIC ANALYSIS IN THE MACHINE
TRANSLATION**

Speciality: 5701.01 – Applied Linguistics

5704.01 – The theory of the language

Field of science: philology – linguistics

Applicant: **Gulnara Natig Jafarova**

Baku –2021

The work was performed in the department of the “Computational Linguistics” at Azerbaijan National Academy of Sciences.

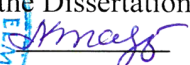
Scientific supervisor: prof., Doctor of Philological Sciences
Kamila Abdulla Valiyeva

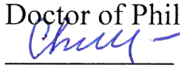
Official opponents: prof., Doctor of Technical Sciences
Alshan Nariman Hasanov

Doctor of Philosophy on Philology,
assoc.prof.
Aygun Shahnazar Mammadova


prof., Doctor of Philological Sciences
Fikret Fetish Jahangirov

Dissertation council – ED 1.06 of the Supreme Attestation Commission under the President of the Republic of Azerbaijan operating at the Institute of Linguistics named after Nasimi, Azerbaijan National Academy of Sciences.

Chairman of the Dissertation council: academician
 **Mohsun Zellabdin Naghisoylu**

Scientific secretary of the Dissertation council:
Doctor of Philosophy on Philology, assoc.prof.
 **Sevinj Yusif Mammadova**

Chairman of the scientific seminar:


Bayram Ganimat Ibrahimov

INTRODUCTION

The actuality and the usage rate of the research work.

Machine Translation is one of the issues being in front of the Computational Linguistics. Generally, Machine Translation is the translation process from one language (source language) to the other language (target language) by using the computer systems without human's help or by the help of human. The field of Machine Translation (MT) has passed certain way of development, some kind of issues have been researched and in this regard, the linguistic terms related to the new technological processes have been discovered.

The basic problems were related to the development of theoretical linguistic models in the Computational Linguistics formed still from the 60 years of XX century. But then the object of the research has been directed to the machine learning, editing of text besides with the applying of statistical methods by enlarging more. *The automatic editing of the texts, automatic morphological, syntactic analysis and synthesis has been researched, more successes have been got in this field. But besides with this, according to the usage degree, if we look through, we can note that, the automatic semantic analysis of the text being one of the special important issues in the Computational Linguistics hasn't been attracted to the research, yet¹.*

The introduced dissertation has been dedicated to the research of the semantic analysis in the MT and its content has very great actuality. We must remember that, the semantic analysis carries out after the morphological and syntactic analysis in MT. But to realize this isn't so easy work. To compile the semantic dictionaries, as well as, homonym, antonym, synonym dictionaries, thesaurus dictionaries is needed for accomplishing the semantic analysis as the linguistic base in the MT. In the dissertation work, all of these issues are in the attention. Except this, the issues of morphological analysis, syntactic analysis are addressed in the process of research. Only the role of

¹ Vəliyeva, K.A. Kompüter dilçiliyinin müasir istiqamətləri // – Bakı: İnformasiya cəmiyyəti problemləri, – 2016. №2, – s. 98-1021

syntactic relations is great in the semantic relations, morphological analysis, in the definition of meaning relation among the root and morpheme, in the sequence of word and word combinations in the sentence during syntactic analysis. Here the semantic relations also play special role besides with the syntactic relations.

Except this, the role and importance of lexical semantics in the MT system during semantic analysis is attracted in the dissertation, the issues of semantic fields and meaning relations are touched in the MT, as well as, the componential analysis of the word is endeavoured. In a word, let's point out this that, firstly, the componential analysis of the words is necessary for realizing the semantic analysis. All of these sayings are conditioned the actuality of the content.

Object and subject of the research. The object of the research is the semantic analysis of scientific-technical texts of the Azerbaijani and English languages in the MT system.

The subject of the research is defining the semantic relations connected the words to each other in the sentences, texts and their components of the Azerbaijani and English languages and preparing the formal models of them.

The aim and the tasks of the research. The main aim of the research work is to define the position of line of the words in the texts for the meaning relations and meaning connection on the basis of the materials of the Azerbaijani and English languages by the way of the semantic analysis. To accomplish the following duties have been conserved for achieving to the aim putting forward:

1. to reveal the semantic models of the words and word combinations for the MT;
2. to solve the semantic relation among the sentence couples related with each other for the grammatical feature;
3. to define the coincidence of which of some possible comments of the revealed part more (the semantic polysemy belongs to here)in the discourse;
4. to appoint the linguistic meanings of the words by the way of semantic analysis in the context of MT, to explain the logical, lexical and grammatical relations among them;

5. to reveal the meanings of components of the words by the way of semantic analysis in the MT;
6. to open the monosemy and polysemy of the words by the semantic analysis;
7. to study the functions of the words or word combinations of the Azerbaijani language inside the sentence or separately, the performed duties;
8. to clarify the ways of translation, expressive means of the words or word combinations of the Azerbaijani language to the English language;
9. to disclosure the inner meanings of the words or word combinations;
10. to research the structure, formal signs of the words or word combinations, to define the appropriate and different features among languages.

The methods of the research. The comparative, descriptive, comparison-confrontation, modeling methods have been mainly used in the dissertation.

The basic provisions giving to the defence: The following provisions related to the semantic analysis in the MT are introduced to the defence:

1. To consider the fact of being different of the meaning levels (the levels of semantic abstraction) of different predicative units is necessary for describing the semantic structure of sentence.
2. To find the lexical meaning of the word is needed before the semantics of the word during the semantic analysis in the MT.
3. To divide the words to the semantic fields according to their meanings is necessary for simplifying the process of semantic analysis in the MT.
4. The lexical units of the language must be grouped under one meaning in the semantic fields. The polysemy must be solved, too here.
5. The meaning colour of the words must be revealed by the context of the words during the process of semantic analysis in the MT.

6. The componential analysis of the words is also necessary during the semantic analysis in the MT.

7. The semantic dictionaries, thesaurus dictionaries must be compiled for the semantic analysis in the MT.

The scientific novelty of the research. The Natural Language Processing (NLP) is one of the most actual issues of period. The level of semantic analysis of the MT system has special importance for the language processing. The accomplishing of semantic analysis in the MT system has been researched in this dissertation firstly. The words have been divided into the semantic fields according to their meanings, have been analyzed for their components and the semantics of the words has been defined by performing from lexical meaning during the process of semantic analysis in the MT system. Here, the tables have been set by separating the words to the groups according to their meanings; too, the relations among them have been expressed. “Latent Semantic Analysis” and the “MapReduce” program have been commended to the research work. Polysemy being one of the most important problems has been researched, the semantic analysis of the polysemantic words has been realized, and the tables defining the grammatical, lexical and semantic description of the words in the sentence and showing the appropriate formula of them have been set in the level of semantic analysis in the MT system. Taking into consideration all of these, the notes listed above can be considered the scientific novelty of the research work.

The theoretical and practical importance of the research. The theoretical importance of the research work consists of this that, the research of the words and word combinations and attracting to the dissertation is applied in the machine learning of the language, in the MT system and it also defines the possibilities of the language system. The process of creation of the words and word combinations, the attracting of them to the research by analyzing from the etymological, typological and semantic feature of them creates a base for learning the degree of fertility in the language. Besides, the polysemy of the words has been analyzed in the sentence beside with the meanings using separately by analyzing from semantic feature in the research work, too.

And the practical importance of the research consists of this that, the results got from the dissertation can be used in the machine translation, in the development of dictionaries, in the special courses belonging to linguistics, in the process of education. Also the doctoral students, the candidates, the high school students can be used from these materials as the useful means.

The approbation and the applying of the work. The basic theoretical provisions and content of the research have been reflected in the lectures of the author introduced to the republic and international conferences, also in more than twenty articles and thesis (generally 29) published in the influential scientific publications in our country and foreign countries.

The name of the organization where the dissertation has been accomplished. The dissertation was accomplished in the “Department of Computational Linguistics” at Azerbaijan National Academy of Sciences named after Nasimi.

The structure of the research work. The dissertation consists of the introduction, three chapters, conclusion and the list of literature.

The volume of the structural sections of dissertation separately and the general volume with the sign. The introduction of dissertation consists of 4 pages, the first chapter is 32 pages, the second chapter is 53 pages, the third chapter is 39 pages, the conclusion is 2 pages, the list of literature section is 22 pages and the list of abbreviations is 1 page. The dissertation consists of a total of 159 pages, consists of 212 221 symbols.

THE BASIC CONTENT OF THE WORK

The actuality of content and the usage rate of the research work is based, the object and the subject of research, the aim and the tasks, the method and ways of the research, the basic provisions giving to the defence are defined, the information about the scientific novelty of the research, the theoretical and practical importance of the work, the approbation and the applying of the work, the name of the organization where the dissertation has been accomplished, the volume of the structural sections of dissertation separately and the

general volume with the sign is given in the part of “**Introduction**” of the dissertation.

The first chapter of the research is entitled as “**The scientific theoretical issues of the Machine Translation**”. The scientific theoretical issues are touched more here. Let’s note that, the demand to the translation of language creates from the need being to the exchange of information and the development of the language relations in the modern period. In such situation, the machine or automatic translation is used as the substitutive method. This process seeing very easy from the first view faces with great practice and difficulties. If it is about the literary translation, one of the obstacles that we meet is the translation of appropriate aphorism, proverbs, the metaphors in the phraseological expressions, in a word, subtle semantic signs here. It is known that, to create formal models of giving factors is very difficult now. But there isn’t the need to formalize the scientific-technical translation. By only approaching from this direction in our subjected dissertation, we aimed the semantic analysis on the texts in the MT. And what is machine translation?

Machine translation (MT) – is the automatic process of the translation works form one language (source language) to the other language (target language) or languages completely or partially by the computer. *Translation, in its full generality, is a difficult, fascinating, and intensely human endeavour, as rich as any other area of human creativity*². Before everything, the translated text (or part of speech) must be understood, perceived whether in the translation by human or translation by machine (in oral or written translation). Human cannot translate the information without perceiving this accepted by the feelings as (ears, eyes, mimes and etc.). *And machine must also accept, “perceive” the text (speech) as*

² Jurafsky, D., *Speech and Language Processing.* / D.Jurafsky, H.M.James. – New Jersey: Prentice Hall, Englewood Cliffs, – 2000, – p.797.

information by the technical means, and then it must pass to the translation level³.

The **first section** of the **first chapter** is entitled as “**The machine translation problem in linguistics and applying in the Azerbaijani linguistics**”.

The translation has been possible from one language to the other language by the prepared electron devices for unlocking the password of secret codes of the enemy in the war years. After the publication of famous “Translation” memorandum by Warren Weaver in 1949, the intensive research works have been applied in this field. *The first practical translation (from Russian to English) by computers has been realized in the USA in 1954. From that year, the researches have been started in this field in the Post Soviet Union, the translation samples have been got from the French and English languages to Russian language in Moscow in 1956*⁴.

The researchers of the Tatar language as F.A.Dreizin⁵, P.K.Bayramova⁶ have researched the syntactic analysis in the MT, the specialists of the Uzbek language as N.A.Khalitova⁷, Kh.F.Iskhskova⁸ have researched the morphological analysis from Uzbek to Russian, over the synthesis Yakubova⁹ (Uzbek language),

³ Mahmudov, M.Ə. Süni intellekt maşın tərcüməsi sistemlərində // Tətbiqi dilçiliyin müasir problemləri. Beynəlxalq elmi konfransın materialları, – Bakı: – 27-28 noyabr, – 2012, – Mütərcim, – s.61-63.

⁴ Mahmudov, M.Ə. Kompüter Dilçiliyi / M.Ə.Mahmudov. – Bakı: Elm və Təhsil, – 2013, – s.98–99.

⁵ Дрейзин, Ф.А. Принцип синтаксического анализа татарской фразы. / Ф.А.Дрейзин, Р.С.Рашитов – Москва: Машинный перевод, – 1961, – 180 с.

⁶ Байрамова, П.К. Переводы на татарский язык. Ний: АКД. / Л., 1966.

⁷ Халитова, Н.А. Машинный перевод с татарского языка на русский (в кн. Проблемы туркологии и истории российского востоковедения) / Н.А.Халитова. – Казань: 1960. – 235 с.

⁸ Исхакова, Х.Ф. Исследования по формальной морфологии тюркских языков (по содержанию татарского литературного языка в переписке с узбекским и узбекским языками): / автореферат диссертация кандидата филологических наук. / – Москва, 1972. – 22 с.

⁹ Исхакова, Х.Ф. Исследования по формальной морфологии тюркских языков (по содержанию татарского литературного языка в переписке с узбекским и узбекским языками): / автореферат диссертация кандидата филологических наук.

Otkupshikova¹⁰ (Mongolian language), G.O.Yefremov¹¹ (Chuvash language) have done researches in the 60-70th years.

Using from the mathematical methods in the Computational linguistics has been realized by the publishing of the book named as “Mathematical linguistics in study” written by M.S.Garayeva in the 60th years in Azerbaijan¹².

After the publication of the monography named as “Mathematical linguistics” written by A.A.Akhundov in 1979, the interest has been increased more¹³.

In 2004, “The frequency dictionary of the Azerbaijani newspaper language” has been published¹⁴. In 2010, “The frequency dictionary of the Azerbaijani language” has been printed¹⁵. In our linguistics, starting to the research of the historical monuments, even the written monuments by the computers have stimulated to the development of alphabet-frequency dictionary and reverse alphabetical list on the basis of the printed text of “Book of Dede Korkut” (compilers: F.Zeynalov, S.Alizade) published in 1988 year¹⁶. Besides this, the statistical analysis of M.Fuzuli’s poetry

наук. / – Москва, 1972. – 22 с.

¹⁰ Откупщикова, М.И. Один из возможных способов формулирования передовой морфологии // – Ленинград: Материал Математическая лингвистика и машинное обучение, – 1963, – с. 61-65.

¹¹ Ефремов, Г.О. К вопросу о машинном переводе текста с русского языка на чувашский. // – Москва: Учен, записки (Чуваш, гос. пед. ин-т им. И. Я. Яковлева), – 1963, – вып.15, – с.79-88.

¹² Караева, М.С. Математическая лингвистика в обучении (К обоснованию проблемы – обучение иностранным языкам по моделям). Азербайджанское Государственное Издательство Учебно-Педагогической литературы / М.С. Караева. – Баку: – 1964. – с.7.

¹³ Axundov, A.A. Riyazi dilçilik. Xülasələr və təfriqələr. Dərs vəsaiti / A.A. Axundov. – Bakı: ADU, – 1979. – s.12.

¹⁴ Vəliyeva, K.A. Azərbaycan qəzet dilinin tezlik lüğəti. / K.A. Vəliyeva, M. Ə. Mahmudov, V.Y. Pines – Bakı: Nurlan, – 2004, – 264 s.

¹⁵ Mahmudov, M.Ə. Azərbaycan dilinin tezlik lüğəti (söz kökləri). / M.Ə. Mahmudov, Ə.B. Fətullayev, S. Məmmədova [və b.] – Bakı: Elm, – с.1. – 2010, – 464 s.

¹⁶ Vəliyeva K.A. “Kitabi-Dədə Qorqud” dastanlarının statistik təhlili / K.A.Vəliyeva, M.Ə.Mahmudov, V.Y.Pines. – Bakı: Elm, – 1999, – 248 s.

being our classical heritage on that period has been realized. Of course, addressing to the classical literature, the researching of stylistic features especially in the authors' works plays special role in the defining of the author of the unknown authored works by comparing the indicators of the frequency of usage.

Except this, the article named as "About some features of the applying of structural and mathematical linguistic methods to the Turkic languages" written by A.A.Akhundov published in 2012 has stimulated to the development of mathematical linguistics, specially¹⁷.

Generally, a large amount of works for this direction have been accomplished in the last years in the Azerbaijani linguistics. For example, we can note the works named as "Electron calculating machines in linguistics" by V.Y.Pines and M.A.Mahmudov¹⁸, "Frequency dictionary of the Azerbaijani newspaper language" by K.A.Valiyeva, M.A.Mahmudov, V.Y.Pines, J.Rahmanov¹⁹, "The transitional dialects of the Azerbaijani language" by A.H.Valiyev²⁰, "Modelling the structure of the Azerbaijani verb forms in the problem of automatic dictionary" by V.Y.Pines²¹, "Development of a system of formal-morphological analysis of the Turkic word forms (on the basis of the materials of the Azerbaijani language" by M.A.Mahmudov²², "The automatic syntactic analysis and synthesis of the text" by K.A.Valiyeva²³, "The development of formal models

¹⁷ Axundov, A.A. Seçilmiş əsərləri: [2 cildə] / A.A.Axundov. – Bakı: Maarif, – c.2. – 2012. – s.172–178.

¹⁸ Pines, V.Y. Elektron hesablayıcı maşınlar dilçilikdə. İcmal məlumatı / V.Y.Pines, M.Ə.Mahmudov – Bakı: AZETETMİ, – 1977, – 36 s.

¹⁹ Vəliyeva, K.A. Azərbaycan qəzet dilinin tezlik lüğəti. / K.A. Vəliyeva, M. Ə. Mahmudov, V.Y. Pines – Bakı: Nurlan, – 2004, – 264 s.

²⁰ Vəliyev, A.H. Azərbaycan dilinin keçid şivələri: / filologiya üzrə elmlər doktoru dis. / – Bakı, 1974. – 400 s.

²¹ Пинес, В.Я. Моделирование структуры азербайджанских глагольных форм в связи с проблемой автоматического словаря: /автореферат диссертация кандидата филологических наук. /–Баку, 1970. –19с.

²² Махмудов, М.А. Разработка системы формального морфологического анализа тюркской словоформы (на материале азербайджанского языка: / диссертация кандидата филологических наук. / – Баку, «Элм», – 1982. – 150 с.

²³ Vəliyeva, K.A. Mətnin avtomatik sintaktik təhlili və sintezi / K.A.Vəliyeva. –

of the Azerbaijani language and the compiling of linguistic processor on their basis” by Z.M.Amirov²⁴, “The compiling and applying of numbering-modelling method for the Azerbaijani-English machine translation system” by A.B.Fatullayev²⁵, “The compiling formal grammar of the “Limited Azerbaijani language” as the component of knowledge basis of deductive machine” by A.M.Khalili²⁶, “Determining the optimal structure of an automatic dictionary and machine translation system” by Z.Y.Guliyeva²⁷, “The algorithmic solution of polysemy of verb in the English-Azerbaijani machine translation system” by A.A.Aliyev²⁸, “Formal description of the synthesis of word”²⁹, “Automatic editing of the texts”³⁰, “The actual issues of the Computational linguistics”³¹, “Applying of probability-statistical methods in the Azerbaijani linguistics”, “The development of information basis of the Azerbaijani language” by K.A.Vəliyeva. In addition, let’s note that, some novelties have been done in the field of MT system of the Azerbaijani language nowadays.

Bakı: Örnək, – 1996, – 158 s.

²⁴ Əmirov, Z.M. Azərbaycan dilinin formal modellərinin yaradılması və onların əsasında linqvistik prosessorun qurulması: / texnika üzrə fəlsəfə doktoru dis. avtoreferatı. / – Bakı, 2006. – 19 s.

²⁵ Fətullayev, Ə.B. Azərbaycan-İngilis məşin tərcüməsi sistemi üçün rəqəmsal modelləşdirmə metodunun işlənilib hazırlanması və tətbiqi: / texnika üzrə fəlsəfə doktoru dis. avtoreferatı. / – Bakı, Kibernetika İnstitutu, 2006. – 19 s.

²⁶ Xəlili, Ə.M. Deduktiv məşinin biliklər bazasının tərkib hissəsi kimi “Məhdud Azərbaycan dili”nin formal qrammatikasının işlənilib-hazırlanması: / filologiya üzrə fəlsəfə doktoru dis. avtoreferatı. / – Bakı, 2009. – s.12.

²⁷ Кулиева, З.Ю. Определение оптимальной структуры автоматического словаря и системы машинного перевода: / автореферат диссертация кандидата филологических наук. / – Баку, 2011. – 46 с.

²⁸ Əliyev, Ə.A. İngiliscə-Azərbaycanca məşin tərcüməsi sistemində felin çoxmənalığının alqoritmik həlli: / filologiya üzrə fəlsəfə doktoru dis. avtoreferatı. / – Bakı, 2012. – 20 s.

²⁹ Велиева, К.А. Формальное описание синтеза слова: /Авторыф.дис. канд. наук./ – Москва, 1971, – 20с.

³⁰ Vəliyeva, K.A. Mətnlərin avtomatik redaktəsi / K.A.Vəliyeva, M.H.Məmmədova – Bakı: “Azərbaycan Universiteti” nəşriyyatı, – 2003, – 80 s.

³¹ Vəliyeva, K.A. Kompüter dilçiliyinin aktual məsələləri // – Bakı: Dilçilik İnstitutunun Əsərləri, – 2015. – s.126-132.

We want to note the development of “Dilmanc” MT system proudly which its presentation has been held and caused to the great sensation in 2005³². Dilmanc – is the first MT system in Azerbaijan.

The great information about the formal approach – modelling in the language being one of different methods of mathematical methods in the Computational linguistics has been given by N.Rajabova in Azerbaijan³³.

In 2013, the monography named as “Computational linguistics” by M.A.Mahmudov, in 2018, the other monography named as “Turkic National Corpus” has been introduced to the readers³⁴. The book named as “The theoretical problems of Machine Translation” by M.M.Mirzaliyeva and K.A.Valiyeva³⁵ has been published in 2016, the monography named as “Introduction to Computational linguistics” by K.A.Valiyeva has been printed in 2018³⁶.

The II section of the first chapter of the dissertation is entitled as “**The levels of analysis to the semantic analysis and translation methods in the Machine Translation**”. The surveys show that, the history of MT surrounds the following levels:

The period which the “Translation” Memorandum has been celebrated under the ruling by Warren Weaver in 1947 is considered the beginning of the first level.

In 1961, the first international conference named as “Applied linguistics and Machine Translation” has been arranged and at first time, the term of “Computational linguistics” has been used here. And only for this, the II level is known the period of the creation of “Computational linguistics”.

³² Fətullayev, Ə.B. Azərbaycan-İngilis maşın tərcüməsi sistemi üçün rəqəmsal modelləşdirmə metodunun işlənilib hazırlanması və tətbiqi: / texnika üzrə fəlsəfə doktoru dis. avtoreferatı. / – Bakı, Kibernetika İnstitutu, 2006. – 19 s.

³³ Rəcəbova, N.Ə. Qrammatik səviyyədə dilin struktur modelləri // – Bakı: 2007

³⁴ Mahmudov, M.Ə. Kompüter Dilçiliyi / M.Ə.Mahmudov. – Bakı: Elm və Təhsil, – 2013, – 356 s. / Mahmudov, M.Ə. Türk dillərinin milli korpusları / M.Ə.Mahmudov. – Bakı: Elm və təhsil, – 2018, – 392 s.

³⁵ Mirzəliyeva, M.M. Maşın tərcüməsinin nəzəri problemləri. / M.M.Mirzəliyeva, K.A.Vəliyeva – Bakı: Elm və Təhsil, – 2016, – 119 s.

³⁶ Vəliyeva, K.A. Kompüter dilçiliyinə giriş / K.A.Vəliyeva. – Bakı: Elm və Təhsil, – 2018, – 384 s.

The III level is considered from 1966 year. In that year, the government of USA has created the committee named as ALPAC (Automatic Language Processing Advisory Committee) for the development of the field of MT. This committee has been called the advisory committee of automatic language processing and increased its work more in the next periods. Even the development of the base of NLP is also related to the name of this committee in the modern period.

Some methods have been created from the first level to the IV level. Here include the Rule-based Machine Translation (RBMT), Directed Machine Translation (DMT), Transfer based Machine Translation (TBMT), Interlingual Machine Translation (IMT), Corpus based Machine Translation (CMT), Hybrid Machine Translation (HMT). From the IV level, it means from 1980 year, the Example-based Machine Translation has been created.

In the V level, Example-based Machine Translation has passed o the other system – Statistical Machine Translation (SMT) from 1990 year.

In its time, some methods of machine translation have been created:

1. Rule-based Machine Translation (RBMT);
2. Corpus based Machine Translation (CMT);
3. Hybrid Machine Translation (HMT)

We have given large information about these methods in the dissertation.

The **first semi section** of the **II semi-chapter** of the **first chapter** of the dissertation is entitled as “**Levels of morphological and syntactical analysis in the Machine Translation and applying in the Azerbaijani linguistics**”. Let’s note that, the morphological and syntactic analysis must be applied till semantic analysis absolutely. For this reason, it would be expedient to give short information about the morphological analysis of text.

Morphological analysis is the most important level of MT system. The morphological analysis is analyzed the structure of the words. Except this, *the morphological analysis gives the possibility to automate the processes of collection of information, accumulation,*

*searching, keeping and passing of information in the information searching system*³⁷. Aim is to separate the word forms of sentence, in a word, the chain of morphemes to the differential parts and to define the grammatical features belonging to them by automatic way in the morphological analysis.

The level of syntactic analysis is accomplished after the morphological analysis. In the level of syntactic analysis, the sentence is chosen and the word, word combination or expressions inside the sentence are analyzed according to formal signs. *“Sentence has the system of special formal features. Sentence itself is used in two meanings: 1) sentence – the sequence of the words, 2) sentence – as the set of the words for expressing the structure of syntactic relations”*³⁸. During the syntactic analysis, the word-form has been accepted as the minimal unit formally in the sentence.

Special attention has also been directed to the MT, fundamental works have been realized in this direction in the Azerbaijani language including to the Turkic languages. So that, much information about the scientific works, novelties accomplished, researched, applied in the Azerbaijani language has given in the first semi-chapter of the first chapter of the dissertation. But in this semi-chapter, the levels of morphological and syntactic analysis of the MT in the Azerbaijani linguistics are touched.

K.A.Valiyeva has researched the morphological analysis largely in her monography named as “Automatic editing of the texts”³⁹.

During syntactic analysis, the syntactic relations among the word-forms arranging the sentence are defined; the members of

³⁷ Белоногов, Г. Г. Автоматизация процессов накопления, поиска и обобщения информации. / Г. Г. Белоногов, А.П. Новоселов, – Москва: «Наука», – 1979. – с.126.

³⁸ Əliyeva, N.Ə. Struktur sintaksisdə sintaktik əlaqələr problemi: / filologiya üzrə fəlsəfə doktoru dis. avtoreferatı. / – Bakı, 2017. AMEA Nəsimi adına Dilçilik İnstitutu, – s.33.

³⁹ Махмудов, М.А. Система автоматической переработки тюркского текста на лексико-морфологическом уровне / М.А. Махмудов. – Баку: ЭЛМ, – 1991, – с.132. / Vəliyeva, K.A. Mətnlərin avtomatik redaktəsi / K.A.Vəliyeva, M.H.Məmmədova – Bakı: “Azərbaycan Universiteti” nəşriyyatı, – 2003, – s.12.

sentence and the relation of coordination are researched. The main aim is the separation of existed sentence to the elements truly or the creation of the sentence which its elements are true in this level. The word groups are investigated only in this level, the basis of sentence is set by the combining of them.

M.A.Mahmudov has set the algorithm of automatic syntactic analysis of the Azerbaijani language in his monography named as “Computational linguistics”⁴⁰.

The III semi-chapter of the **first chapter** of the dissertation is entitled as “**Lexical Semantics in the Machine Translation**”. We must note about the semantic feature of language unit that, the lexical semantics being one of the field of semantics deals with the meaning of word.

While dealing with the semantics, Y.Seyidov expresses that, *the semantic meaning is the meaning which is acquired in the situation of speech or text and it depends on lexical meaning. The semantic meaning is the meaning amenities appearing by different ways or means from lexical meaning. Though the semantic meaning is related to concept, it is not the free expression of it. For this, we use the expression of “lexical semantics”. Semantic meaning is inside of the lexical meaning. It is not separate from lexical meaning; it cannot live apart from it*⁴¹. Agreeing with this idea, too, we want to note that, the lexical semantics is analyzing the relations among the meanings of words. During the formal analysis, the revealing of the meanings of the components of word, giving and revealing of the monosemantic or polysemantic meaning of word must found its reflection in the lexical semantics. But the research scientist has noted in his work that, *semantic meaning is inside the lexical meaning*. By disproving this thought, we want to note that, the expression of “semantic meaning” is false. The word “semantics” is called as “being related to meaning, having the meaning”. Here, if we say “the semantics of the word is inside the lexical meaning”, it

⁴⁰ Mahmudov, M.Ə. Kompüter Dilçiliyi / M.Ə.Mahmudov. – Bakı: Elm və Təhsil, – 2013, – s.251–296.

⁴¹ Seyidov, Y.M. Seçilmiş əsərləri: [15 cildə] / Y.M.Seyidov. – Bakı: BDU, – c.9. – 2010, – s.210.

can be truer.

Lexical semantics studies how arranging of the lexicon of language, how relating the meanings of lexical units with each other oppositely. The main aim of the lexical semantics is to set the models for the structure of the lexicon of language by separating the types of relation to the categories among the words. Lexical semantics is based to the linguistic meaning. By considering the notes, we can say that, the role of lexical semantics is important in the implementation of semantic analysis in the Machine Translation system.

The II chapter of dissertation is entitled as “**The level of semantic analysis in the Machine Translation**”. It is known that, the formal description of semantics is necessary in the description of the complex process of language, especially in the issues related with the semantic analysis. In linguistics, formal semantics is the revealing of relations of certain meaning among the expressions existing in natural language and setting of mathematical models of them. It means, “Formal semantics” reflects the notion of semantic analysis in the MT system. We say the semantic analysis of formal systems, so what is semantics?

“Semantics” has taken from the word “semantikos” (σημαντικός) being the ancient Greek origin word, it is called “being related to meaning, having the meaning”. Semantics is the term being related with the meaning. Therefore, “*the meaning is the part of the language; semantics is one of the sections of the linguistic science dealing with the meaning*”⁴².

The first semi-chapter of the **II chapter** of the dissertation is entitled as “**The semantic analysis of the word in the Machine Translation**”. One of the first from the most important works is the compiling of general semantic dictionaries, thesaurus dictionaries for semantic analysis. It is the most part of the work. Then the semantic relations must be defined. *One of most effective ways of revealing of*

⁴² Palmer, F.R. Semantics [II edition] / F.R.Palmer. – Cambridge: Cambridge University Press, – 1983, – p.1.

*the semantic relations is the method of setting lexical models*⁴³.

Researcher A.Sokirko has noted specially that, *semantic unit is one of central concepts of semantic theory. It arranges the foundation of semantic base*⁴⁴. Agreeing with A.Sokirko's thought, we must note that, semantic unit – is the object of textual semantics.

The first section of the first semi-chapter of this chapter of dissertation is entitled as **“Semantic analysis and structure of the polysemantic words in the Machine Translation”**. *The new meaning appearing from the first meaning of any word is called the colour of its meaning. The words expressing the different meaning colours of this concept are called polysemantic words*⁴⁵. “Polysemantics” is the Greek word for its origin, it has appeared from the combinations of words “çox” (many) (polos) and “işarə” (sign) (sema), the meaning is “çoxmənəlı” (polysemantic)⁴⁶.

A.Gurbanov has given such definition to the polysemantic words: *“In addition to its main meaning, expressing some meanings of the words is called the polysemy of the words by related to using in different sentences”*⁴⁷. Polysemantics has developed more in most of the natural languages. As the sample to this, we can show the English, French and many of the languages. For example, let's look through the meanings of the word “tail” in the modern English language:

1) the back of the body of a bird, animal or fish that can be moved - quyruq (quş, heyvan və ya balığın bədəninin hərəkət edən

⁴³ Lyons, J. Introduction to Theoretical Linguistics / J.Lyons. – Cambridge: Cambridge University Press, – 1968. – p.428. / Мочалова, А.В. Алгоритм семантического анализа текста, основанный на базовых семантических шаблонах с удалением // – Scientific and Technical Journal of Information Technologies, Mechanics and Optics, – 2014. №5 (93), – с.128-129.

⁴⁴ Сокирко, А.В. Семантические словари в автоматической обработке текста (по материалам системы ДИАЛИНГ): / диссертация кандидата технических наук. / – Москва, 2001. – 120 с.

⁴⁵ Sadıqova, A. Mürəkkəb və çoxmənəlı sözlər, fərqi, müqayisəli təhlili və əhəmiyyəti // – Bakı: Filologiya məsələləri, – 2013. №4, – s. 118.

⁴⁶ <https://en.wikipedia.org/wiki/polysemy>

⁴⁷ Qurbanov, A.M. Müasir Azərbaycan ədəbi dili: [2 cildə] / A.M.Qurbanov. – Bakı: Nurlan, – c.1. – 2003. – s.317–318.

arxa hissəsi); for example: “The dog yelped when the guy stepped on his tail” (Uşaq itin quyruğunu tapdalayanda it hürməyə başladı).

2) the back/end of sth - bir şeyin axırı, sonluğu (məsələn, the back of a cart (arabanın arxası)); for example: “He fell backwards, must have hit his head on the *tail of the car*”⁴⁸ (O, arxası üstə düşdü və başı maşının baqaj hissəsinə dəydi).

3) the meaning of plait belong to people, the back part of the hair from the top to the bottom - hörükcük (insan saçının arxa kökündən ucunadək forma ilə yığılmış hissəsi); for example: “Pulling her hair into a pony tail, she secured it with a blue ribbon”⁴⁹ (O, saçına hörük vuraraq, mavi lentlə bağladı”).

4) the meaning of the line of people that stand back by back - sıra, növbə mənasında, for example: “There was a tail in front of the bakery” (Çörək dükanının qarşısında quyruq vardı) - burada sıra mənasında işlənilir.

5) the back part of a plane - təyyarənin quyruğu; for example: “Unprotected at the tail of the aircraft they led precarious lives” (Təyyarənin göyərtəsinin arxa hissəsində onlar təhlükəli vəziyyətdə müdafiəsiz qaldılar).

6) the back side of a coin that does not have a picture of the head of a person on it - metal pulun gerb olmayan arxa tərəfi; for example: “Each coin has two sides: the tail and the head”(Hər pulun iki üzü var: arxa və ön üzü).

7) a long jacket divided at the back below the waist into two pieces that become narrower at the bottom, worn by men at very formal events - frak (kişilərə məxsus geyim növü, arxası iki hissəyə bölünmüş, arxa ucu dar uzun kostyum); for example: “In the last periods, many of the men liked to wear the tail” (Keçmişdə kişilərin əksəriyyəti frak geyinməyi sevirdi).

8) a person who is sent to follow sbd secretly and find out information about where that person goes, what they do - təqibçi, təqib edən, bir kimsəni izləyən; for example: “Yesterday someone tailed him” (Dünən onu kimsə təqib edirdi).

⁴⁸ <https://context.reverso.net/перевод/английский-русский>

⁴⁹ <https://sentence.yourdictionary.com/tail>

9) the back part of sth/sb - dal, daldakı, geridəki; for example: “The tail of the house was destroyed by someone” (Evin arxasını kimsə dağıtmışdır).

Nine meanings of the word “tail” have been already cleared to us. If we take notice attentively, we can see that all of these meanings are closed to each other, the expression of “arxa, geri, dal” (back) reflects itself in the explaining and usage of every of them. But only one of these meanings is the first meaning, the others are derivatives from this. In the ancient English language, the first meaning of the word “tail” has been given as “*tægl, tægel*”, the word has been used in the meaning of “uzunsov, nazik tük” (long, thin hair). Let’s note this here that, being the first meaning which of the polysemantic words and the derivative meaning which of the polysemantic words in their meaning colours and the historical source of that word must be researched from etymology of them.

In the Azerbaijani linguistics polysemantics is also chosen with its complexity in the MT system. Generally, we should emphasize that, the meaning system of any polysemantic word develops by the adding of new meanings to the old meanings during centuries. In the Azerbaijani linguistics, the polysemantic words are come across in the noun, adjective and verb more being the notional parts of speech.

Let’s note that, the polysemantic word expresses only one lexeme. *In case of being the relation among two or more different lexemes in homonymy, the polysemy is the feature of one lexeme*⁵⁰.

Two types of approach influence to the analysis of semantic structure of the polysemantic words: diachronic approach and synchronic approach.

In polysemantics, beside the word is retaining the previous meaning, it can acquire two or more new meanings itself in the same time, it gives the new meanings in different situations during the diachronic approach. For example, there are many meanings of the

⁵⁰ Lyons, J. Language and Linguistics. An Introduction / J.Lyons. – UK: Cambridge University Press, – 1981, –p.68.

word “*stol*” (*table*) in the modern English language:

1. A piece of the furniture (table, chair and etc.) Mebelin bir hissəsi (stol, stul və s.);
2. The people sitting at the table - Stol arxasında oturan camaat;
3. The meal put on the table - Stol üzərində olan yemək;
4. Board - Sərlövəhə;
5. Slabs of stone (with words written on them) – Daş plitələr (üzərində döyülərək sözlər yazılmış plitələr);
6. Words cut into slabs of stone – Daş plitələrə yazılmış nəsihətamiz sözlər (Bibliyadan);
7. an orderly arrangement of facts or figures – Cədvəl (faktlar və ya fiqurları olan);
8. Part of a machine-tool – Dəzgah;
9. A level of area – Yayla (coğrafi termin);
10. Plateau – Məcməyi.

All meanings of the word “*Stol*” (*table*) are existed in the modern English language. The first meaning of this word is the word “*tabule*”, as “*taxta və ya daşın dəyirmi plitəsi*” using in the ancient English language. In the ancient English language, the word “*tabule*” has taken from the word “*tabule*” being in the Latin language. The other meanings of the word “*tabule*” (*table*) has taken from the first meaning and has been created as diachronically. We can get such result that, the existing of two types of the meaning emerges in the diachronic approach:

1. The first meaning of the word;
2. The second meaning of the word (derivative meaning later).

The synchronic approach is understood as the existing of different meanings of the same word in the appointed historical period of development in polysemantics.

It is considered that, *the issues of disappearing polysemy, refusing of homonymy whether in the lexical-morphological or syntactic levels are the work of semantic analysis of system completely*⁵¹. While one polysemantic word is used in the sentence,

⁵¹ Белоногов, Г.Г. Алгоритм многоступенчатого морфологического анализа русских слов / Г.Г. Белоногов, Ю.П.Калинин, М.Ф.Поздняк [и др.], – НТИ: – 1983. №1 (2), – с.8.

firstly to find the using of it from which context in the automatic analysis of this sentence is needed. For this, the grammatical and lexical analysis must be done before them. For example, let’s make the table by using such a sentence (Table 2.1.1.3):

Table 2.1.1.3.
Table defining the grammatical, lexical and meaning
description of word or expressions
“Gecə-gündüz səhər açır, gecə isə laçəklərini yumur”

Word or expression	grammatical analysis	description of meaning
Gecə-gündüz		
001gecə-gündüz	noun, common noun, singular, subject	the flower name
002gecə-gündüz	noun, common noun, compound, singular, subject	time expressing evening and morning
.....
...		
səhər		
001səhər	adverb, adverb of time, adverbial modifier of time	time of sunrise
002səhər	noun, common noun, singular, subject	time of sunrise
003səhər	noun, proper noun, singular, subject	personal name
.....
açır		
001açır	verb, indicative mood of the verb, present tense, III person singular	to open smth.
002açır	verb, reflexive verb, present tense	opening of smth (flower, weather and etc.)
003açır	verb, indicative mood of the verb, present tense, III person singular	open the account, number for smbd.
004açır	verb, indicative mood of the verb, present tense, III person singular	open a debate

Continuation of the table 2.1.1.3.

005açır	verb, indicative mood of the verb, present tense, III person singular	turn on the tap
006açır	verb, indicative mood of the verb, present tense, III person singular	clear the way
007açır	verb, indicative mood of the verb, present tense, III person singular	unveil the monument
008açır	verb, indicative mood of the verb, present tense, III person singular	unravel the mystery
009açır	verb, indicative mood of the verb, present tense, III person singular	open the dress, button, belt and etc.
010açır	verb, indicative mood of the verb, present tense, III person singular	open the hook, package
011açır	verb, indicative mood of the verb, present tense, III person singular	untie the knot
012açır	verb, indicative mood of the verb, present tense, III person singular	roll out the dough
013açır	verb, indicative mood of the verb, present tense, III person singular	lay the table
.....
geçə		
001geçə	adverb, adverb of time, adverbial modifier of time	time of sunset
002geçə	noun, common noun, singular, subject	time of sunset
.....

Continuation of the table 2.1.1.3.

isə		
001isə	secondary parts of speech, conjunction	the method using for comparing two facts, events and etc.
.....
ləçəklərini		
001ləçəklərini	noun, common noun, plural, object	the colourful, smelling part of flower
002ləçəklərini	noun, common noun, plural, object	scarf worn over the head by the women (kerchief, shawl)
.....
yumur		
001yumur	verb, reflective verb, present tense	closing of smth.(flower)
002yumur	verb, indicative mood of the verb, present tense	close the eyes and etc.

In this level, we note some alternatives for choosing the coincidence which of them to the context more during the morphological analysis of the sentence. Thus, the word “*gəcə-gündüz*” is used as the function of subject at the beginning of sentence and after this we chose the adverb of time “*səhər*” as the adverbial modifier by coinciding to the word order of the Azerbaijani language. Till there in the analysis, we can use the proper noun “*Səhər*” being the personal name. But in the next levels of the analysis we appear impossibility of this. Then the verb “*açır*” is used in the sentence and with this we can see the concluding of the sentence. Here the verb “*açır*” has been given in some meanings. But we need to find the true polysemantic word that, it would coincide to that context. The sentence that we used for analysis is the complex sentence. And for this, we continue the analysis after the quotation mark comma “,”. The second part of the analysis of the complex sentence starts with the word “*gəcə*” and here it performs in

the function of the adverb of time. Then the conjunction “*isə*” is used and it accomplishes the function of confrontation among two sentences. The noun “*Ləçəklərini*” is used after the conjunction of “*isə*” and it fulfills in the function of object. Already we affirmed our thought that has said before by the usage of the word “*ləçəklərini*”. At the end of the sentence, the verb “*yumur*” carries the function of predicate. Then everything is known to us. This sentence is about the flower, the description of the flower, not about the personal name. After realizing the grammatical, lexical, morphological and syntactic analysis, we can carry the semantic analysis of the sentence.

The **second semi-chapter** of the **second chapter** of dissertation is entitled as “**The semantic analysis of word combination and sentence in the Machine Translation**”. *The notion of word combination is the relation of different words with each other using in the same place in text combining by the syntactical way*⁵².

The word combination creates from the meaning and grammatical relation of two or more words. Thus, 3 basic conditions must be followed for the appearing of the word combination:

- 1.The semantic relation among the components of word combinations;
- 2.The syntactic relation among the components;
- 3.The situation, position, order, rule of order.

In the Azerbaijani linguistics, the researchers as G.Sh.Kazimov, M.Huseynzade, Y.M.Seyidov, A.Abdullayev, A.Hasanov, N.Ch.Valiyeva and the others have been researched the word combination⁵³.

⁵² Sidorov, A. G. Automatic Syntactic Analysis for Detection of Word Combinations / A.G.Sidorov, A.Gelbukh, S.Y.Han [et al.] // Computational Linguistics and Intelligent Text Processing, – Mexico: – 2004, N2945, – p. 240-244.

⁵³ Kazimov, Q.Ş. İsmi birləşmələr və onların tərəfləri arasında sintaktik əlaqələr // – Bakı: Dilçilik məsələləri (məqalələr toplusu), Elm, – 2000, – s.51. / Hüseynzadə, M. Təyini söz birləşmələri. Azərbaycan dilinin qrammatikası / M.Hüseynzadə. – Bakı: EA, – 1959, – s. 60-67. / Abdullayev, Ə.Z. Müasir Azərbaycan dili. IV hissə, Sintaksis. / Ə.Z. Abdullayev, Y.M.Seyidov, A.Q. Həsənov – Bakı: Şərq-Qərb, –

K.A.Valiyeva has researched the formal description of ordering of word combinations in detail over the machine translation system in the Azerbaijani language⁵⁴. In this dissertation, we also tried to realize the researching of word combinations for semantic analysis in detail by referencing to her research.

In this section, the most difficult level is the researching of the problems of semantic analysis of the sentence being impossible for formalization in the MT system.

A.Horak also notes that, *the basic possibility of semantic analysis is to focus to the searching of true descriptions of the meaning of sentence*⁵⁵. According to this, to set the models of the sentence is one of the basic conditions for the lexical meaning and structure. Let's note that, the structural model of the sentence and lexical construction has been compiled by M.S.Garayeva in the Azerbaijani linguistics at first time. She has separated the sentences to the divisions on the basis of structural scheme and she has set the construction of sentence according to lexical approach⁵⁶.

Related to the semantic analysis of the sentence in the MT system, to touch two problems is needed. The first is to solve the semantic relation among the sentence doubles communicated to each other for grammatical feature. The second is to define which of some possible comments of the determined part is coincided more (here include the semantic polysemy or polysemantics).

The word can be translated differently during the analysis in the text, some of them can be inappropriate and unnecessary to the context. For example, let's give the sentence using with the word

2007. – s.49. / Vəliyeva, N.Ç. Müxtəlif sistemli dillərdə feili birləşmələrin tipoloji təhlili / N.Ç.Vəliyeva. – Bakı: Avropa, – 2008, – 169s.

⁵⁴ Vəliyeva, K.A. Kompüter dilçiliyinə giriş / K.A.Vəliyeva. – Bakı: Elm və Təhsil, – 2018, – s.333.

⁵⁵ Horak, A. The normal translation algorithm in transparent intensional logic for Czech / A.Horak. – Brno: Czech Republic, Faculty of Informatics, Masaryk University, – 2002, – p.3.

⁵⁶ Караева, М.С. Математическая лингвистика в обучении (К обоснованию проблемы – обучение иностранным языкам по моделям). Азербайджанское Государственное Издательство Учебно-Педагогической литературы / М.С. Караева. – Баку: – 1964. – с.7.

“külfət”:

“Namaz kişinin külfətiyəm” – deyə, qarı könülsüz cavab verdi” (M.Huseyn)⁵⁷. If we translate this sentence into English language, we can substitute the translation of the word “külfət” with the word “wife”. Whereas, there are some meanings in the translation of the word “külfət” into English language, it means, some words can be used here, for example:

1)household (oğul-uşaq,arvad-uşaq mənasında);

2)family (ailə,nəsil, soy, tayfa mənasında);

3)mate (yoldaş, külfət mənasında);

4)wife (arvad, həyat yoldaşı, zövcə, əyal mənasında).

But according to the context, the last meaning form this word meaning is coincided more, it means, “Namaz kişinin zövcəsi, arvadı” is appropriate.

Let’s look through the other sample using with the word “külfət”, for example: “Pinəçilik edib yeddi baş külfətin ruzisini bir tövr ilə qazanıram” (S.S.Axundov)⁵⁸. In this sentence, we must substitute the word “külfət” with the word “family”, not with the word “wife” into English by depending on the context.

While translating the sentence, the above meanings can be expressed as the translation of the word “külfət”. But it is not true to say that, every of them can be used during the analysis of the sentence. If we look through the context, we can define that the word “külfət” has been used here in the meaning of “Pinəçilik edib yeddi baş ailənin ruzisini bir təhər qazanmaq”. Besides, the word “külfət” can be used in the other meaning by depending on the context. For example, the word “külfət” can also be used in the meaning of “zəhmət, əziyyət, qayğı” in the abstract meaning. Let’s show such sample: “Onu külfət basıb”.

As noted by A.Khalili, “the form of the sentence depends on

⁵⁷ Azərbaycan dilinin izahlı lüğəti [Dörd cildə] II cild / – Bakı: Şərq-Qərb, – 2006, – s.776.

⁵⁸ Yenə orada, – s.776.

*the content and can be in different forms*⁵⁹.

*The semantic analysis is defined for appointing the meaning of the unit word for every word in the sentence. The basic issue is to create the large amount of semantic word groups in the semantic analysis*⁶⁰.

K.A.Valiyeva has compiled the algorithmic scheme for checking the structure of sentence automatically in the MT system⁶¹.

The III semi-chapter of the II chapter of dissertation is entitled as “The semantic analysis of the text in the Machine Translation”.

The semantic analysis of the text describes the relations among the linguistic elements in the MT. The semantic relation of linguistic elements defines the relation of arguments to each other. Non-existing of semantic relation in text is caused to the losing of the meaning passing from one part to the other part of the text.

The first section of the III semi-chapter of this chapter of dissertation is entitled as “The methods of semantic analysis of the text in the Machine translation”. Different methods have been set for showing the meaning, conception and content in the semantic analysis of the text in the MT. These methods help to define the problems during the semantic relation among the elements of the text by depending on various features of the text. Here we have shown some methods offered by some linguists for analyzing different semantic features of the text in the MT system.

The III chapter of the dissertation is entitled as “The program software of the semantic analysis in the Machine Translation”. In this chapter, the issues of the semantic fields and

⁵⁹ Xəlili, Ə.M. Deduktiv məşının biliklər bazasının tərkib hissəsi kimi “Məhdud Azərbaycan dili”nin formal qrammatikasının işlənilib-hazırlanması: / filologiya üzrə fəlsəfə doktoru dis. avtoreferatı. / – Bakı, 2009. – s.12.

⁶⁰ Simmons, R.F. A semantic analyser for english sentences / R.F.Simmons, J.F.Burger // Mechanical Translation and Computational Linguistics, –Santa Monica, California: – 1968. vol.11, nos.1 and 2, March and June, – p.1.

⁶¹ Vəliyeva, K.A. Mətnlərin avtomatik redaktəsi / K.A.Vəliyeva, M.H.Məmmədova – Bakı: “Azərbaycan Universiteti” nəşriyyatı, – 2003, – 80 s.

meaning relations are touched, except this; the componential analysis of the word during semantic analysis is researched. As the program software, Latent Semantic Analysis and MapReduce programs are looked through.

The first semi-chapter of the III chapter of dissertation is entitled as **“The semantic fields and meaning relations in the Machine Translation”**. The theory of semantic fields reflects the grouping of the lexical units of language under one meaning in the MT.

The notion of existing of certain semantic groups, including to one or some such groups of the language units arranges the base of the theory of semantic field⁶².

One of the main principles of semantic analysis is to define the belonging of word to which word groups, to which semantic field, being or not being polysemantic word in the MT. To define the relations of any word including to the semantic field and being or not being related to the other words is very important in this direction. One of the main works is the accomplishing of semantic analysis in the same time for defining the meaning of the word during the process of semantic analysis in the MT. Except this; the semantic field must be compiled better by clarifying the arranged semantic relation with the other closed or different meaningful notions. In the semantic field, the role of one meaning and the relation with the other word meanings must be clarified in the result of the semantic analysis.

It is known to us that, the word groups being closed to each other for the meaning arrange the semantic field. The words including to the semantic field are grouped according to the closeness of the expressed thing (event) and notions not being related to each other from the etymological feature.

The meaning relations have special role in the MT. By depending on the characters of relations among the words, different –

⁶² Hüseynova, G. Tat dilinin lüğət tərkibinin leksik-semantik təhlili / G.Hüseynova. – Bakı: Elm və təhsil, – 2014, – s.25.

lexical, semantic, lexical-semantic, thematic groupings of the words are possible. We noted some types of the meaning in the research work. About it much information has been given in the dissertation.

The II semi-chapter of the **III chapter** of the dissertation is entitled as **“The componential analysis of the word during the process of semantic analysis in the MT”**.

The componential analysis is one the modern directions of semantic analysis. The componential analysis separates the meaning to the smallest components and is analyzed them. From the first researchers ordering the method of componential analysis, we can give the names of some American linguists as F.G.Lounsbery, V.H.Gudenau and the other. So that, related to the componential analysis, the researchers have done enough researches. We can note some of them as J.Fodor, J.Katz, E.A.Nida, U.Weinreich, I.V.Arnold, R.S.Ginzburg, E.M.Mednikova, O.N.Seliverstova, I.A.Sternin and the others.

From the Russian researchers as L.A.Volkovs, O.V.Kondrashova in the analysis of different group nouns, J.P.Sokolovskaya, L.V.Lavrova, T.S.Shepina in the analysis of meaning of the adjectives and L.M.Vasilyev, A.M.Kuznetsov in the analysis of the meaning of the verbs have been used the method of componential analysis.

L.M.Vasilyev expresses that, *while saying the componential analysis, the method of semantic analysis realizing with the considering of the minimal meaning units arranging different meanings and these meanings of the words is usually understood*⁶³.

According to Y.S.Stepanov, *the choosing of the meanings is getting by the method of the opposite attitudes in the different meaning of them, in addition, by the way of opposite comparison of the meanings with the meanings of other words in the componential analysis*⁶⁴.

R.S.Ginzburg notes that, *the expressing of the concrete words*

⁶³ Васильев, Л.М. Методы семантического анализа // Исследования по семантике. – Уфа: – 1975. вып. I, – с. 158.

⁶⁴ Степанов, Ю.С., Гипотеза в современной лингвистике. / Ю.С.Степанов, О.Н. Селиверстова – Москва: Наука, – 1980, – с.287–288.

*consists of the separation of essence features of expressed object in the componential analysis*⁶⁵.

According to the thought of L.I.Shakhova, *the choosing of the meanings is getting by the way of abstracting of the characteristic features of the facts naming by the words during the componential analysis*⁶⁶.

J.P.Sokolovskaya notes in her research work that, *the componential analysis performs as the reflection of the reality of semantic meanings of the language*⁶⁷.

The III semi-chapter of the **III chapter** of dissertation is entitled as **““Latent Semantic Analysis” and “MapReduce” program in the Machine Translation”**. Special programs are also compiled for accomplishing of semantic analysis in the MT. To this, we can show “Latent Semantic Analysis” such as sample. The compiling of this program is known to us from the world linguistics. In the translation to the English language, it is called as “Gizli semantik təhlil” (GST) “Latent Semantic Analysis” (LSA). Sometimes it is called as “gizli semantik kodlaşdırma” (GSK) (Latent Semantic Indexing) (LSI). This analysis is called as “Probabilistic Latent Semantic Analysis” in the English language. LSA defines the closed meaningful words existing in the similar texts. The applied directions of LSA are these:

- 1.To compare the documents (grouping information);
- 2.To emerge similar information among the languages after the analysis of the translated documents;
- 3.To find the relations among the terms (synonym or polysemantic words);
- 4.To reveal the appropriate documents;
- 5.To define the most true coincidence or similarity among the

⁶⁵ Гинзбург, Р.С. Значение слова и методика компонентного анализа // Иностранные языки в школе, – 1978, №5, – с.21.

⁶⁶ Шахова, Л.И. Компонентный анализ семантики абстрактных существительных (Семантика языка и текста.) / Л.И.Шахова. – Кировоград: – 1984. – с. 50-51.

⁶⁷ Соколовская, Ж.П. Модел семантических отношений в лексике / Ж.П. Соколовская. – Киев: – 1986, –с.43.

smallest groups of terms by the semantical way in the frame of corpus;

6.To enlarge the applied directions of Machine Translation system;

7.To analyze the word relation in the text corpus.

LSA is the process analyzing the relation among the information appearing together. LSA is used largely in the processing of the automated information. But it gets much time in the accomplishing of the analysis in the large amount of data base and often makes difficulties. Related to this, MapReduce program has been created. MapReduce is applied in distributed form for solving the analysis of large scale texts using in the computer programs.

The IV semi-chapter of the III chapter of the dissertation is entitled as “Dictionaries during the semantic analysis of the word in the Machine Translation (Thesaurus dictionaries)”. The main base is the development of the dictionaries for the semantic analysis of text. We tried to make research belonging to the thesaurus dictionaries in our dissertation because of connected with the semantic analysis of it. Thus, the thesaurus is the dictionary of meaning of the words that grouped for meaning, not for the alphabet. The second function of the thesaurus is the explanatory description of semantic relations reflecting the communications of knowledge field. Therefore, *we can approach to the thesaurus dictionary as the semantic model of the translation language*⁶⁸.

*The thesaurus gives the opportunity to the systematization of all lexical fond of the language paradigmatically. In these dictionaries, the defining of semantic attitudes of the words plays an important role*⁶⁹.

Certain works related to the development of thesaurus have done, information retrieval thesaurus for building has been prepared

⁶⁸ Vəliyeva, K.A. Mətnlərin avtomatik redaktəsi / K.A.Vəliyeva, M.H.Məmmədova – Bakı: “Azərbaycan Universiteti” nəşriyyatı, – 2003, – s.38–39.

⁶⁹ Чижаковский, В.А. Тезаурус в системах автоматической переработки текста. / В.А.Чижаковский, Л.Н. Беляева – Кишинев: Штиинца, – 1983. – с.75.

*in the Azerbaijani linguistics*⁷⁰.

*As the semantic model of researched scientific field, thesaurus serves to the revealing of the concrete meanings of the language units (word, word combinations) inside the text*⁷¹.

The development and improvement of new generation computers more gradually gives the opportunity to the explanation and setting of semantic analysis again in the MT system.

Realization of semantic analysis is the basic factor influencing to the development of information communication technologies strictly in the MT system.

The following main conclusions were got in the research work:

The processing of the text using in the natural language is one of the actual issues in the contemporary period. Nowadays the machine translation has become a reality of new information technologies. Only for this, we have chosen the level of semantic analysis playing a special role in the machine translation, we have touched to the research of syntactic relations among the words on its basis in our dissertation. We can summarize our thoughts related to the study of models of analysis in the machine translation system:

1. The semantics of the word is defined by the lexical meaning of it during the semantic analysis in the MT. So that, semantics depends on the lexical meaning and it is inside the lexical meaning.

2. In the process of semantic analysis, the words are divided into the semantic fields for their meanings in the MT. Let's note that, the systematic character is related to the "semantic fields" in the lexicology and semantics of the language. The lexical units of the

⁷⁰ Мамедова, М.Г. Автоматизированный отбор лексики в информационно-поисковый тезаурус на основе анализа терминологических словарей: / автореферат диссертация кандидата филологических наук. / – Москва, 1984. – 25 с. / Мамедова, М.Г., Пинес В.Я. Принципы построения информационно-поискового тезауруса по строительству // Тез.докл. Семинара-совещ. «Республиканская система научно-технической информации и опыт организации информационного обеспечения народного хозяйства», – Баку: – 1989, – с. 39-41.

⁷¹ Mahmudov, M.Ə. Türk dillərinin milli korpusları / M.Ə.Mahmudov. – Bakı: Elm və təhsil, – 2018, – s.277–278.

language are grouped under one meaning in the semantic fields. It means, we have collected the lexical units belonging to special field by basing to the workable principle of the analysis in the semantic field, we have commented the meaning relations among the words by researching their meanings. In a word, the words have been classified according to the types of meaning.

3. In the same time, the words are analyzed according to their components during the process of semantic analysis in the MT. We must note that, the semantic components define how the members belonging to one group are related to each other or how they are differed from the other by depending on the surrounding of one lexeme in the componential analysis.

4. The result we got is that, the compiling of semantic dictionaries, thesaurus dictionaries having an important role in the semantic analysis is the necessary issue in the MT system. Considering to this, we have prepared some samples belonging to these dictionaries and demonstrated in the dissertation.

5. We think so that, to realize the etymological analysis is possible beside with the semantic analysis of the word in the MT.

6. Let's consider that, the polysemy must be solved during the process of semantic analysis in the MT. We must note that, the main source of polysemantics consists of the existing of changings in the semantic structure of the word. So that, the performing of the polysemantic word from which context firstly is found during the automatic analysis of sentence. Primarily, the grammatical and lexical analysis is accomplished for this.

The following thesis and articles relating to the subject of the dissertation have been published:

1. Evolutionary Process of the Computational Linguistics in Azerbaijan // – USA: The USA Journal of Applied Sciences. CIBUNET (ORT/Publishing), – 2015. № 6, – p. 12-14

2. Statistik məşın tərcüməsində paralel korpuslar (türk dilləri materialları əsasında) // – Bakı: Azərbaycan Texniki Universiteti, “Elmi-əsərlər” jurnalı, – 2016. Cild 1, №3, – İSSN 1815-1779, – AzTU mətbəə, – s. 106-109.

3. Məşın tərcüməsində polisemantik sözlərin semantik təhlili

və quruluşu // “Azərbaycan filologiyası: İnkişafın yeni mərhələsi” mövzusunda Respublika elmi konfransının materialları, – Bakı: “Bakı Universiteti” nəşriyyat, – 2 noyabr – 2016, – s. 143-149.

4. Maşın tərcüməsində semantik təhlilin orta q ünsiyyət dilinin yaradılmasına təsiri və əhəmiyyəti // Türkoloji Elmi-Mədəni Hərəkətdə Ortaq Dəyərlər və Yeni Çağırışlar” mövzusunda Beynəlxalq Konfrans, – Bakı: Elm və təhsil, – 14-15 noyabr – 2016, – s. 249-252.

5. Semantic context of the word in the machine translation // Müasir Dilçiliyin Aktual Problemləri” Beynəlxalq Elmi Konfransın materialları, – Sumqayıt Dövlət Universiteti, Sumqayıt: – 24-25 noyabr, – 2016. – p.267-269.

6. Maşın tərcüməsində təhlil mərhələləri // – Bakı: Tədqiqlər, – 2017. №5, – İSN – 5 – 8066 – 0933 – 2, – s. 131-139.

7. The role of computer software in teaching the Azerbaijani language to the various systematic language representatives // “Ağamusa Axundov və Azərbaycan Filologiyası” Beynəlxalq elmi konfransın materialları, – Bakı: – 24-25 aprel, – 2017. – İSN-978-9952-514-04-9, – “Elm” nəşriyyat, – p. 425-429.

8. Mənanın təhlilində semantik komponentlərin rolu // Doktorantların və gənc tədqiqatçıların XXI Respublika elmi konfransı, – Bakı: Bakı Dövlət Universiteti, – 24-25 oktyabr – 2017, – s. 118-119.

9. Leksik semantikada sözün komponentlərinə görə təhlili // – Bakı: Dilçilik İnstitutunun əsərləri, – 2018. №1, – “Elm və təhsil” nəşriyyat, – s. 180-186.

10. Maşın tərcüməsində çoxmənalılığın semantik təhlili və quruluşuna baxış // – Bakı: Azərbaycan Milli Elmlər Akademiyası, Gənc Alim və Mütəxəssislər Şurası, “Gənc Tədqiqatçı” Elmi-praktiki jurnal, – 2018. IV cild, №2, – “Elm” RNPM-nin mətbəəsi, – ISSN 2409-4838, – s. 121-128.

11. Maşın tərcüməsinə nəzəri baxış // – Bakı: Filologiya məsələləri, – 2018. №4, – “Elm və təhsil” nəşriyyat, –ISSN 2224-9257, – s. 147-157.

12. Maşın tərcüməsində sözün semantik təhlilinin yerinə yetirilməsində etimoloji təhlilin rolu // – Bakı: Filologiya məsələləri,

– 2018. №16, – “Elm və təhsil” nəşriyyat, – ISSN 2224-9257, – s. 140-146.

13. Maşın tərcüməsində sözün semantik təhlili // – Bakı: Dil və ədəbiyyat, – 2018. IX cild, №2, – Dillər Universiteti, – s. 73-80.

14. Azərbaycan dilçiliyində maşın tərcüməsinin morfoloji və sintaktik təhlil mərhələləri // AMEA, Nəsimi adına Dilçilik İnstitutu, “Müstəqillik illərində üslubların inkişafı və dilin lüğət tərkibinin zənginləşmə istiqamətləri” mövzusunda Respublika elmi konfransının materialları, – Bakı: – 25-26 dekabr, – 2018, – “ELM”, – s.167-171.

15. Maşın tərcüməsində sözün semantik təhlilində lüğətlər (Tezaurus lüğətləri) // – Bakı: AMEA Filologiya və Sənətsünaslıq, – 2019. №1, – ISSN 2663-4368, – UOT 81, – s. 193-196.

16. The semantic analysis methods in the computational linguistics // – Bakı: AMEA, Tədqiqlər, – 2019. №2, – s. 5-13.

17. Научно-Теоретические Вопросы Машинного Перевода и Его Применения в Азербайджанском Языке // – Казань: Казанская наука, – 2019. № 12, – Издательство Рашин Сайнс, – с. 152–155.

18. Türk dünyasında maşın tərcüməsi probleminə bir baxış // Bakı Avrasiya Universiteti, – VI Beynəlxalq Türk Dünyası Araşdırmaları Simpoziumu, – Bakı: – 13-15 iyun, – 2019, – Cild 3, – “Elm və təhsil”, – s. 99-104.

19. Afad Qurbanovun tətbiqi dilçiliyin tədrisində rolu // Tanınmış dilçi alim, türkoloq və ictimai xadim, professor, Azərbaycan Milli Elmlər Akademiyasının (AMEA) müxbir üzvü Afad Qurbanovun anadan olmasının 90 illiyi münasibətilə respublika elmi konfransı, – Bakı: – 24 may, – 2019, – s.138-142.

20. Machine Translation in the Turkic Languages (mainly in the Azerbaijani language) // The VII International Conference on Computer Processing of Turkic Languages “TurkLang 2019”, – Simferopol, Crimea, Russian Federation: – 2019. – October 3-5.

21. Maşın tərcüməsində sözün semantik təhlili // – Bakı: AMEA, – “Türkologiya” Beynəlxalq elmi jurnalı, – 2020. № 2, – 36–45.

22. The research of the Machine Translation problems and

applying in the Azerbaijani Linguistics // XXVII International Scientific-Practical conference «EurasiaScience», – Moscow, Russia: Research and Publishing Center, –15 February, – 2020, – p.225–229.

23. Maşın tərcüməsində lüğətlər və onların işlənmə prinsipləri // AMEA-nın Nəsimi adına Dilçilik İnstitutunun yaradılmasının 75 illiyinə həsr olunan “Dilçiliyin aktual problemləri” adlı beynəlxalq elmi konfrans, – Bakı: noyabr, – 2020.

24. Semantic Relation among The Words and Word Models During The Semantic Analysis of the Text in the MT // II International Scientific Conference of Young Scientists and Specialists, Multidisciplinary approaches in solving modern problems of fundamental and applied sciences, – Azerbaijan, Baku: – 03-06 March, – 2020, – p.428–430.



The defense will be held on the 29th of June in 2021 at 13⁰⁰ at the meeting of the Dissertation council – ED 1.06 of Supreme Attestation Commission under the President of the Republic of Azerbaijan operating at Azerbaijan National Academy of Sciences, the Institute of Linguistics named after Nasimi.

Address: Baku, AZ 1143, The avenue H.Javid 115, V floor, ANAS, The Institute of Linguistics named after Nasimi.

Dissertation is accessible at the Library of Azerbaijan National Academy of Sciences, The Institute of Linguistics named after Nasimi.

Electronic versions of dissertation and its abstract are available on the official website of the Institute of Linguistics named after Nasimi, Azerbaijan National Academy of Sciences.

Abstract was sent to the required addresses on the 27th of May in “2021”

Signed for print:25.05.2021
Paper format: 60x84.16¹
Volume: 35, 508
Number of hard copies: 20