

UOT-811.512.162

Səhifə:91-95

<https://doi.org/10.59849/2663-8967.2023.2.91>

*Nigar Nağıyeva*  
*AMEA, Nəsimi adına Dilçilik İnstitutu*  
*nagiyevanigar8@gmail.com*

## KOMPÜTER DİLÇİLİYİ TERMINOLOGİYASININ MƏNBƏLƏRİ

### XÜLASƏ

Kompüter dilçiliyi sahəsində dilin tədqiqi və onun fəaliyyət proseslərinin modelləşdirilməsi üçün müxtəlif kompüter proqramları və təbii dil məlumatlarının tərtibi və emalı məqsədilə məlumat texnologiyaları hazırlanır və istifadə olunur. Bu elmi və təbii sahədə xüsusi adların sayının sürətlə artması müvafiq terminologiyanın həddlərinin dinamik və bəzən nəzarətsiz genişlənməsinə kömək edir. Məqalədə də “kompüter dilçiliyinin terminologiyası” ifadəsinə aid edilən sözləri analiz edilmişdir. Kompüter dilçiliyinin fənlərarası sahəsinin peşə dili dilçilik, riyaziyyat, informatika və süni intellektin terminoloji vahidlərini ehtiva edir. Kompüter dilçiliyi anlayışına yaxın olan “təbii dilin avtomatik işlənməsi” terminoloji ifadəsidir ki, bu ifadə öz mənasında dil məlumatlarını emal etmək üçün kompüterlərin istifadə olunduğu bütün sahələri əhatə edir.

Terminologiya məsələlərinin öyrənilməsinə həsr olunmuş çalışmalardan, kompüter dilçiliyi və təbii dilin avtomatik emalı mövzuları üzrə əsərlərdən, leksikoqrafik və ensiklopedik mənbələrdən istifadə etməklə aparılmış tədqiqat nəticəsində kompüter dilçiliyi terminologiyasının kompüter texnologiyasından istifadə əsasında təbii dil əsasında modelləşdirməsi ilə əlaqəli anlayışları kodlaşdırmaq üçün nəzərdə tutulmuş xüsusi dil vahidləri sistemi kimi təyin edilməsi təklif edilmişdir.

*Açar sözlər:* kompüter dilçiliyi, terminologiya, təbii dil, kodlaşdırma, vahid.

**Giriş:** Kompüter dilçiliyi terminologiyasının formalaşması, kompüterin müxtəlif ictimai sahələrdə işləndiyi və təbii dil məlumatlarının emalı üçün kompüter texnologiyalarının təcrübədə geniş istifadə edildiyi bir dövrdə baş vermişdir. Hazırda bütün fəaliyyət sahələrinə və insanların gündəlik həyatına kompüter avadanlıqları, internet şəbəkəsi, “ağıllı” və çoxfunksiyalı mobil telefonlar, yeni məlumat-axtarış sistemləri, kompüter tərcüməsi, səsli köməkçilər, virtual həmsöhbətlər və s. sürətlə daxil olmaqda davam edir. Bu proseslər getdikcə artan və geniş istifadə olunan müvafiq terminləri xarakterizə edən terminlərin yaranmasına səbəb olur. Tədqiqatın aktuallığı ondadır ki, kompüter dilçiliyi dilçilik və informatikanın kəsişməsində yerləşən nisbətən yeni və mürəkkəb bir elmdir. Hazırda bu sahənin terminoloji sisteminin öyrənilməsinə həsr olunmuş xüsusi tədqiqatlar nisbətən azdır.

Kompüter dilçiliyi alətləri və texnologiyalarının əhəmiyyətli inkişafı Azərbaycan dilində bu sahədə terminlərin əmələgəlmə üsullarına, quruluşuna və digər xüsusiyyətlərinə görə fərqlənən çoxlu sayda yeni sözlərin meydana gəlməsi ilə nəticələndi. Bu terminlərin, demək olar ki, hamısı ingilis dilindən alınmadır. Bu alınmalar sırasına embedinq, teq, parsinq, brauzer kimi sözləri qeyd edə bilərik. Bəzi terminlər reterminologiya nəticəsində ikinci məna qazanmışlar: məsələn, korpus, konveyer, transformer, vektor və s. Bir sözdən ibarət terminlərin (klaster, rəqəmsallaşdırma) və terminoloji ifadələrin (avtomatik nitq sintezi, səsli köməkçi, avtomatik tamamlama, maşın tərcüməsi) istifadəsi geniş yayılmışdır. Dilçilikdən kənar proseslərin sürətlə və nəzarətsiz şəkildə artması xüsusi kompüter adlarının sistemləşdirilməsi prosesini çətinləşdirərək müvafiq sahənin terminologiya sərhədlərini genişləndirir.

“Kompüter dilçiliyi terminologiyası” anlayışının əhatə dairəsi hələ də dəqiq müəyyən edilməmişdir. Azərbaycan dilinin izahlı lüğətində termin sözü [lat. terminus – hədd, sərhəd] elm, texnika,

incəsənət və s. müəyyən ixtisas sahəsinə aid anlayışı dəqiq ifadə edən söz (və ya söz birləşməsi) kimi təqdim edilmişdir [1, s.278]. A.A.Reformatski yazır ki, terminlər “özlərinin xüsusi təyinatı ilə məhdudlaşan xüsusi sözlərdir. Termin "məhdud mütəxəssislər dairəsi tərəfindən istifadə olunan konseptual aparatın tərkib hissəsidir" [2]. İnsan fəaliyyətinin müəyyən bir sahəsinə aid söz və ifadələr toplusu müvafiq terminoloji sistemi təşkil edir. Dilçilikdə terminologiya anlayışı müxtəlif sərhətlərə malikdir. İ.Q.Barabanovanın fikrincə, elmi terminlərdə vacib olan amil “onun formalaşmasının spesifikliyidir ki, bu da əksər hallarda peşəkar səviyyədə ümumiləşdirilmiş və standartlaşdırılmış daha mücərrəd konseptual xüsusiyyətlərdən ibarətdir” [3].

Mütəxəssislərin fikrincə, kompüter dilçiliyi dilçilik, riyaziyyat, informatika və süni intellektin kəsişməsində yaranmış fənlərarası xüsusi bir elm sahəsidir. Kompüter dilçiliyinin lüğət tərkibinə yuxarıda göstərilən elm sahələrinin terminoloji vahidləri daxildir, məsələn: leksema, söz forması, mətn, dil norması (dilçilik terminləri); qraf, əmsal, statistik model (riyaziyyat terminləri); məlumat bankı, istifadəçi interfeysi, skript (informatika terminləri); neyron şəbəkəsi, transformator, çərçivə (süni intellekt terminləri).

“Kompüter dilçiliyi” termini ingilis dilindəki *computational linguistics* terminindən kalka yolu ilə düzəlmişdir. Onu da qeyd etmək ki, *computational* sözü həm də riyazi sözü kimi də tərcümə edilir. Beləliklə də, yeni meydana çıxan istiqamətin əvvəllər riyazi dilçilik adlandırılması bu səbəbdən olmuşdur. Qeyd etmək ki, bu məsələ “kvantativ dilçilik” anlayışının təyini ilə əlaqələndirilir. Tətbiqi dilçiliyə daxil olan kvantativ dilçilik riyazi statistik metodların tətbiqi ilə dilin müxtəlif səviyyələrində baş verən dəyişiklikləri, onların baş vermə səbəblərini və ümumi mənzərəsini müəyyən etməyə imkan verir. Müasir elmi ədəbiyyatda kvantativ dilçilik termini dil və nitqin öyrənilməsi üçün kəmiyyət və ya statistik təhlil üsullarının əsas vasitə kimi xidmət etdiyi tətbiqi tədqiqat istiqaməti mənasında geniş istifadə olunur.

Nəzərdən keçirilən konsepsiyanın fənlərarası əlaqəsi kompüter dilçiliyində müəyyən məlumatın ötürülməsi, habelə mətnin avtomatik işlənməsi və danışq dilinin tətbiqi problemlərinin həlli məqsədilə dil biliklərinin kompüter modelləşdirilməsi ilə məşğul olan süni intellekt sahəsi kimi çıxış edir.

Kompüter dilçiliyi sahəsi çərçivəsində həll olunan tapşırıqlar sırasına təbii dildə insan-kompüter qarşılıqlı əlaqəsinin baş tutmasına imkan verən ünsiyyət modelləşdirilməsi daxildir. Təbii dildə mətnlərin təhlili və yaradılması proseslərinin kompüter modelləşdirilməsi, səslənən nitqin tanınması və sintezi, həmçinin mətnin və səslənən nitqin qarşılıqlı çevrilməsinin tətbiqi problemlərinin həlli (məsələn, məndəki sözlərin qeyri-müəyyənliyinin həlli, maşın tərcüməsi, danışanın səsi ilə müəyyən edilməsi, dikturun identifikasiyası və s.) əlaqəli istiqamət, adətən, təbii dilin avtomatik emalı termini ilə qeydə alınır. Bu terminoloji ifadə öz mənasına görə linqvistik məlumatların işlənməsi üçün kompüterlərin istifadə olunduğu bütün sahələri əhatə etməlidir. Təcrübədə isə bu termin dialoq sistemlərinin, səsli (virtual) köməkçilərin və çatbotların yaradılması ilə əlaqəli tədqiqatların adlandırılmasında istifadə olunur.

Kompüter dilçiliyinin kifayət qədər inkişaf etmiş digər sahəsi məlumat axtarışını həyata keçirmək, sənədlərin indeksləşdirilməsi, ümumiləşdirilməsi və təsnifatı ilə bağlı problemlərin həlli, habelə sual-cavab sistemlərinin yaradılması üçün kompüter vasitələrindən istifadə olunmasıdır. Təbii dilin avtomatik işlənməsi ilə yanaşı, kompüter dilçiliyinə korpus dilçiliyi, elektron lüğətlərin yaradılması, tezauruslar, ontologiyalar, simvolların avtomatik tanınması, kompüter təhlili və nitq sintezi daxildir. Bu elmi və tətbiqi sahənin ən mühüm sahəsi son illərdə həyata keçirilən maşın tərcüməsidir. Kompüter dilçiliyinin cari tətbiqi vəzifələrinə mətnlərdən məlumat çıxarmaq, onların tonallığını təhlil etmək (emosional rəngləmə) və fikirləri müəyyən etmək daxildir [4]. Bu vəzifələr həmçinin verilənlərin intellektual analizi çərçivəsində də həll olunur.

Riyazi əsasların dilçiliyə nüfuz etməsi bu səbəbdən asılıdır: dilin inkişafı dilin təhlili üçün daha dəqiq və obyektiv metodların tətbiqini tələb edir. Dil materialının sistemləşdirilməsi, ölçülməsi və ümumiləşdirilməsində riyazi üsulların istifadəsi nəticələrinin keyfiyyətli interpretasiyası ilə birlikdə dilin strukturunu anlamağa imkan yaradır. Dilçilik elminin başqa elmlərlə əlaqəsinin qurulması yalnız riyazi dildən istifadə etməklə mümkündür. Dilçilikdə riyazi modellərdən istifadə tam həlli

olmayan intuitiv şəkildə tərtib edilmiş linqvistik problemi alqoritmik həlli olan daha sadə və məntiqi tərtib olunmuş riyazi məsələlərlə əvəz etməyə yönəlmişdir. Riyazi dilçiliyin əsas anlayışları ilkin simvollar məcmuəsi (əlifba, lüğət) və onların ardıcılığının çoxluğu (zəncirlər, söz formalarının çoxluğu, söz birləşmələri) məfhumudur. Riyaziyyat elminin mücərrəd anlayış və üsullarının mənimsənilməsi dilçilik elmi üçün mövcud imkanları genişləndirir və müəyyən qanunauyğunluqların kəşfinə töhfə verir. Dil faktlarını öyrənmək üçün riyaziyyatın müxtəlif sahələrindən istifadə olunur: cəbr, çoxluqlar nəzəriyyəsi, riyazi məntiq, informasiya nəzəriyyəsi, ehtimal nəzəriyyəsi və riyazi statistika. Riyaziyyat sayəsində tədqiqatlar daha dəqiq olur, dilin fəaliyyət şəraitini görmək, dil proseslərinin səbəblərini anlamaq, dilin inkişafını proqnozlaşdırmaq və s. mümkün olur. Kompüter dilçiliyi qeyd olunan riyazi terminlərlə fəaliyyət göstərir: aksiomatik metod, kəmiyyət təhlili, linqvistik-riyazi model, linqvistik obyekt və ya hadisənin riyazi izahı, ümumi linqvistik məcmu, mətnlərin linqvistik (dil) korpusu, tezlik lüğəti və s. Informasiya texnologiyalarının inkişafı və internetin qloballaşması xüsusi terminologiyanın, o cümlədən dilçilik kimi bir elm sahəsinin formalaşmasına və sistemləşdirilməsinə səbəb olmuşdur. S.V.Qrinyova görə, hər hansı bir bilik sahəsinin terminologiyası "xüsusi məqsədlər üçün dillərin semantik nüvəsini təşkil edir. Riyaziyyatın alt dilinin semantik nüvəsini riyazi terminologiya, kompüter alt dilini – termin sistemləri şəklində olan kompüter terminologiyası təşkil edir" [5]. Həm riyazi, həm də kompüter terminləri kompüter dilində fəaliyyət göstərir və kompüter biliklərinin kodlaşdırılması, toplanması və ötürülməsinə xidmət edən xüsusi işarə sistemindən ibarətdir. Riyazi və kompüter terminologiyası təbii dilin özünəməxsus əməliyyat alqoritmləri və inkişaf qanunauyğunluqları olan mürəkkəb alqoritmik sistem kimi kompüter tərəfindən çözülməsi mümkün olan problemləri əhatə edir. Bu terminologiyalar linqvistik baxımdan xüsusi yer tutur. Kompüter terminologiyası iki başlıq altında nəzərdən keçirilə bilər: aparat və proqram təminatı. Bu baxımdan kompüter texnologiyaları ilə bağlı dilə çox sayda terminlər daxil olmuşdur (əməliyyat sistemi, formatlaşdırma, instalyasiya - quraşdırma, tabulyasiya, idarəetmə paneli və s.).

Kompüter terminologiyasının inkişafı ekstralingvistik amillərdən asılıdır. Kompüter terminologiyasının fərqli bir xüsusiyyəti başqa dillərdən əhəmiyyətli miqdarda alınma sözlərin işlənilməsidir (latın, ingilis, rus, fransız və alman dillərindən). Kompüter terminologiyasında dillərarası müdaxilə informasiya texnologiyaları və digər fəaliyyət sahələri arasında fənlərarası əlaqələrin olması ilə izah olunur. Kompüter terminologiyasında iki növ termin var: işarəli terminlər (disk, displey) və fraza terminləri (müjərrəd məlumat növü, ilkin proqram yükləyicisi). Sonuncular nominasiyanın ən məhsuldar vahidləri kimi şərh olunur, onlar birmənalılıqı və ardıcılıqı təmin edir, nominasiyanın tələblərinə adekvat cavab verir və inteqral, mürəkkəb konsepsiyanın bir sıra diferensial xüsusiyyətlərini çatdırır. Məzmun strukturuna görə birmənalı (bayt) və çoxmənalı terminlər (şəbəkə) fərqləndirilir. Semantik nöqtəyi-nəzərdən terminsiz ifadələr (qrafik təsvir) və sabit ifadələr (sistem inzibatçısı) fərqləndirilir. Formal strukturuna görə - termin-sözlər, sadə (səhifə), düzəltmə (keçid), mürəkkəb (video giriş) və mürəkkəb qısaldılmış (İnternet) növlərinə bölünür.

V.P.Danilenko simvol-sözləri qeyd edir. Belə ki, sözügedən sahədə beşdən çox sözdən ibarət ətraflı terminlər var (Hypertext Markup Language - HTML) [6]. Həmçinin terminlər yalnız sözdən və ya abreviaturdan ibarət ola bilər. Bəzi hallarda abreviaturlar sözlərlə birgə işlənir. Məsələn: MGQ generator.

Dildə antonimlərin yaranması, inkişafı və formalaşması onların aid olduqları maddi aləmdəki əşya, hadisə və hərəkətlərin ziddiyyətli olması ilə əlaqədardır [7, s.5]. Bu ziddiyyətlər əksmənalı sözlərlə ifadə olunur. Kompüter dilçiliyinin lüğəti daxilində antonim cütlər, əsasən, tamamlayıcı əlaqələri ifadə edir, bunlar təzadlı növ anlayışları arasında ara terminlərin olmaması ilə xarakterizə olunur: giriş – çıxış təbii dil məlumatları, dar – geniş fonetik transkripsiya, kortəbii – düşünülmüş nitq. Kompüter dilçiliyinin terminoloji sistemində antonim vahidlər komponentlərin sayına görə fərqlənir: 1) tək sözdən ibarət (prefiks – postfiks); 2) iki sözdən ibarət (relevant sənəd – qeyri-relevant sənəd); 3) üç sözdən ibarət (açıq mətn korpusu – qapalı mətn korpusu); 4) ətraflı (mövzu modellərinin daxili keyfiyyət meyarları – mövzu modellərinin xarici keyfiyyəti meyarları).

Kompüter dilçiliyinin antonimik terminləri struktur xüsusiyyətlərinə görə müxtəlif köklü (24,8%) və eyni köklü (75,2%) terminlər arasında fərqləndirilir. Müxtəlif köklü antonimlərə misal

olaraq aşağıdakı cütləri göstərmək olar: düzünə – əksinə maşın tərcüməsi; ontologiyanın yuxarı – aşağı səviyyəsi; qaydanın sol – sağ hissəsi.

**Nəticə.** Yuxarıda göstərilənlərə əsasən, kompüter dilçiliyinin terminologiyası praktiki məsələlərin həlli üçün kompüter vasitələrindən istifadə əsasında təbii dilin öyrənilməsi və modelləşdirilməsi ilə bağlı olan və konseptual, leksik-semantik, söz yaradıcılığı və qrammatik səviyyələrlə əlaqəli terminlər çoxluğudur. Kompüter dilçiliyinin terminologiyası bir tərəfdən yetkin sistem, digər tərəfdən isə daim yenilənən və elm və texnikanın digər sahələrinin terminologiyaları ilə, eləcə də ümumi dillə qarşılıqlı əlaqədə olan açıq sistemdir. Son onilliklərdə hesablama texnologiyası və süni intellekt sahəsindəki irəliləyiş təkcə kompüter dilçiliyində terminlərin sayının sürətlə artmasına deyil, həm də onların əhəmiyyətli transformasiyasına səbəb olmuşdur. Belə şəraitdə onların adekvat terminoloji identifikasiyasını həyata keçirmək olduqca çətin məsələdir. Bu tendensiya bir çox terminologiya üçün xarakterikdir və tədqiq olunan xüsusi adlandırmalar sistemində də ifadə olunur. Aparılan təhlil “kompüter dilçiliyi terminologiyası” anlayışını kompüter texnologiyalarından istifadə əsasında təbii dil modelləşdirməsi ilə əlaqəli anlayışların kodlaşdırılması üçün nəzərdə tutulmuş xüsusi dil vahidləri (leksik və qrammatik) sistemi kimi formalaşdırılmasına imkan verir.

### ƏDƏBİYYAT

1. Azərbaycan dilinin izahlı lüğəti, IV cild, Bakı, Şərq-Qərb, 2006, 712 səh.
2. Реформатский А.А. Термин как член лексической системы языка // Проблемы структурной лингвистики 1967. - М.: Наука, 1968. - С. 103-125.
3. Барабанова И.Г. Когнитивный аспект современной русской экономической терминосистемы. Нальчик, 2006. -20 с.
4. Баранов А.Н. Компьютерная лингвистика//Большая российская энциклопедия. URL: <https://old.bigenc.ru/linguistics/text/2087783> (дата обращения: 12.09.2023).
5. Гринев С.В. Введение в терминоведение. - М.: Московский Лицей, 1993. - 309 с.
6. Даниленко В.П. Русская терминология. Опыт лингвистического описания. М.: Наука, 1977. 246 с.
7. Həsənov H.Ə. Azərbaycan dilinin antonimlər lüğəti. Bakı, Şərq-Qərb, 2007, 144 səh.
8. Mahmudov M.Ə. Kompüter dilçiliyi. Bakı. Elm və Təhsil, 2013, 356 səh.
9. Vəliyeva K.A. Kompüter dilçiliyinə giriş. Bakı. Elm və Təhsil, 2018, 384 səh.
10. Лейчик В.М. Терминоведение: предмет, методы структура. М.: Либроком. 2009. 256 с.
11. Суперанская А.В., Подольская Н.В., Васильева Н. В. Общая терминология. Вопросы теории. М. : Либроком. 2012. 248 с.

*Nigяр Nаgiyeva*

*Источники терминологии компьютерной лингвистики*

*Резюме*

В области компьютерной лингвистики разрабатываются и используются различные компьютерные программы и информационные технологии сбора и обработки данных естественного языка для изучения языка и моделирования процессов его деятельности. Быстрый рост числа специальных названий в этой научно-прикладной области способствует динамичному, а иногда и неконтролируемому расширению границ соответствующей терминологии.

Цель статьи - проанализировать значение слов, относящихся к термину «терминология компьютерной лингвистики». Профессиональный язык междисциплинарной области компьютерной лингвистики включает терминологические единицы лингвистики, математики, информатики и искусственного интеллекта. Близким к понятию компьютерной лингвистики является терминологическое выражение «автоматическая обработка естественного языка», которое по своему смыслу охватывает все области, где компьютеры используются для обработки лингвистических данных.

В результате исследований проведенных с использованием работ, посвященных изучению вопросов теоретической терминологии, работ по компьютерной лингвистике и автоматической обработке естественного языка, лексикографических и энциклопедических источников выяснилось, что терминология компьютерной лингвистики определяется как система специальных языковых единиц, предназначенных для кодирования понятий, связанных с моделированием естественного языка на основе использования компьютерных технологий.

**Ключевые слова:** *компьютерная лингвистика, терминология, естественный язык, кодирование, единица.*

*Nigar Nagiyeva*

*Sources of computational linguistics terminology*

*Summary*

In the field of computational linguistics, various computer programs and information technologies for the collection and processing of natural language data are developed and used for studying the language and modeling its activity processes. The rapid growth of the number of special names in this scientific and applied field contributes to the dynamic and sometimes uncontrollable expansion of the boundaries of the corresponding terminology. The purpose of the article is to analyze the meaning of words related to the term «terminology of computational linguistics». The professional language of the interdisciplinary field of computer linguistics includes terminological units of linguistics, mathematics, computer science and artificial intelligence. Close to the concept of computational linguistics is the terminological expression "automatic processing of natural language", which in its meaning covers all areas where computers are used to process linguistic data. On the basis of conducted studies, using works devoted to the study of issues of theoretical terminology, works on computational linguistics and automatic processing of natural language, lexicographic and encyclopedic sources it is proposed to define the terminology of computational linguistics as a system of special language units intended for coding concepts related to the modeling of natural language based on the use of computer technologies.

**Key words:** *computational linguistics, terminology, natural language, coding, unit.*

**Çapa tövsiyə edən:**

*AMEA, Nəsimi adına Dilçilik İnstitutu*

**Rəyçilər:**

*dosent Rəşim Heydərov*

*dosent Svetlana Novruzova*