

# AZƏRBAYCAN RESPUBLİKASI

*Əlyazması hüququnda*

## **BƏDİİ MƏTN LİŖQVOSTATİSTİK TƏDQİQAT OBYEKTİ KİMİ**

İxtisas: 5704.01 – Dil nəzəriyyəsi

Elm sahəsi: Filologiya

İddiaçı: **Həmidə Əzizağa qızı Xəlilova**

Fəlsəfə doktoru elmi dərəcəsi  
almaq üçün təqdim edilmiş dissertasiyanın

### **AVTOREFERATI**

**BAKI – 2025**

Dissertasiya işi AMEA Nəsimi adına Dilçilik İnstitutunun Süni İntellekt və kompüter dilçiliyi şöbəsində yerinə yetirilmişdir.

Elmi rəhbər: filologiya elmləri doktoru, professor  
**Məsud Əhməd oğlu Mahmudov**

Rəsmi  
opponentlər: filologiya elmləri doktoru, professor  
**İlham Mikayıl oğlu Tahirov**

filologiya üzrə fəlsəfə doktoru, dosent  
**Cavanşir Xankişi oğlu Muradov**

filologiya üzrə fəlsəfə doktoru, dosent  
**Səbinə Əli qızı Süleymanova**

Azərbaycan Respublikasının Prezidenti yanında Ali Attestasiya Komissiyasının Azərbaycan Milli Elmlər Akademiyası Nəsimi adına Dilçilik İnstitutu nəzdində fəaliyyət göstərən ED 1.06 Dissertasiya şurası

Dissertasiya şurasının  
sədri: filologiya elmləri doktoru,  
professor  
**Nadir Balaoğlan oğlu  
Məmmədli**

Dissertasiya şurasının elmi  
katibi: filologiya üzrə fəlsəfə doktoru,  
dosent  
**Sevinc Yusif qızı Məmmədova**

Elmi seminarın  
sədri: filologiya elmləri doktoru,  
professor  
**Bahar Cumay qızı Cəfərova**



## GİRİŞ

Dissertasiya N. Gəncəvi, M.Ə. Sabir, M. Füzuli və A.S. Puşkinin bədii mətnlərinin linqvostatistik və korpus tədqiqinə həsr olunmuşdur. Tədqiqata N.Gəncəvinin “Xəmsə”sinə daxil olan poemalarının Azərbaycan və qismən rus dillərinə tərcümələri, M. Füzulinin Azərbaycan dilində yazdığı “Divan” və poemaları, A.S. Puşkinin əsərləri (“Kapitan qızı” və s.), həmçinin M.Ə. Sabirin iki cildlik “Hophopnamə”sində təqdim olunan satira, qəzəl, məktub, hekayə və felyeton və digər əsərləri daxil edilmişdir.

**Mövzunun aktuallığı və işlənmə dərəcəsi.** Müasir dövrdə korpus dilçiliyi və linqvostatistika elmin ən sürətlə inkişaf edən sahələrindən birini təşkil edir. Bu sahələr nisbətən yeni yaranmasına baxmayaraq, çox qısa müddətdə sürətli inkişaf mərhələsinə qədəm qoymuşdur. İnformasiya texnologiyaları dövründə onların inkişafı geri dönüşü olmayan xarakter daşıyır və bu səbəbdən müstəqil elm sahəsi statusunu qazanmışdır. Korpus dilçiliyi və linqvostatistik metodlar sayəsində tezlik lüğətləri, konkordanslar, mətn korpusları tərtib etmək, mətn massivlərinin müqayisəli təhlillərini aparmaq mümkündür. Eyni zamanda, mətnin orijinal və tərcümə dilləri arasındakı fərqləri aşkar etmək, bədii mətnlərin dil xüsusiyyətlərini müəyyənləşdirmək olar. Linqvostatistik və korpus metodlarının tətbiqi leksik müxtəliflik əmsalını (LMƏ) hesablamanı, tokenlərin sayını müəyyənləşdirməyi və daha bir çox əməliyyatları mümkün edir.

Tədqiqatın aktuallığı həmçinin Azərbaycan dilində mətn korpuslarının yaradılmasının vacibliyi ilə müəyyən olunur, çünki indiyədək Azərbaycan dili mətn korpusları ilə təqdim olunmamışdır.

Dissertasiyanın aktuallığını artıran digər amil seçilmiş materialların xarakteridir. Belə ki, M. Füzuli, M.Ə. Sabir və N. Gəncəvinin bədii mətnləri indiyə qədər korpus və linqvostatistik araşdırmaların obyektı olmamışdır. Materialların seçimi də təsadüfi deyildir. N. Gəncəvi, M. Füzuli və M.Ə. Sabirin şəxsiyyəti və onların yaradıcılıq irsi daim aktual mövzular sırasındadır. Məşhur şair və filosof N. Gəncəvinin “Xəmsə”si həm Azərbaycanda, həm də onun

hüdüdlərindən kənarında hərtərəfli tədqiqat obyektinə çevrilmişdir. Nizami irsinin tədqiqatının əhəmiyyətini Nizami Gəncəvi Mərkəzi ilə Oksford Universiteti arasında AMEA-nın vitse-prezidenti, Moskva Dövlət Universitetinin Bakı filialının rektoru, akademik Nərgiz Paşayevanın təşəbbüsü və iştirakı ilə imzalanmış müqavilə də təsdiqləyir.

Azərbaycan satirik poeziyasının banisi, görkəmli şair M.Ə. Sabirin satiraları bu gün də öz aktuallığını qoruyub saxlayır. Sabir qabaqcıl bir şəxsiyyət idi. Satiralarında qaldırdığı mövzuların mürəkkəbliyi və aktuallığı ilə böyük satirik öz dövrünü xeyli qabaqlamışdır. M.Ə. Sabirin yazı üslubu və dili də tədqiqatımız üçün böyük əhəmiyyət kəsb etmir. O, sadə və aydın yazaraq, canlı xalq danışığı ilə ədəbi dili birləşdirərək poeziyada yeni üslub yaratmışdır. Məhz bu xüsusiyyətinə görə Sabirin poetik dili xalqa asanlıqla başa düşülən idi və onun yaradıcılığının geniş kütlələr arasında yayılmasına şərait yaratmışdır.

Azərbaycan ədəbiyyatının digər mühüm nümayəndəsi – M. Füzuli – görkəmli şair, filosof, mütəfəkkir və tərcüməçidir. Bütün bunlarla yanaşı, M. Füzuli ədəbi Azərbaycan dilinin banisidir ki, bu da tədqiqat üçün mühüm əhəmiyyət daşıyır.

A.S. Puşkin isə müasir rus ədəbi dilinin banisi hesab edilir. O, kitab dili ilə rus xalq nitqini birləşdirməklə milli-ədəbi dili yaratmış, kilsə slavyanizmləri, avropa mənşəli sözləri və xalq danışığı elementlərini sintez etmişdir. Puşkinin yaradıcılıq irsinin əhəmiyyətini müxtəlif sahələrin alim və mütəxəssisləri dəfələrlə vurğulamışdır. Ümummilli lider H. Əliyevin də A.S. Puşkin yaradıcılığına yüksək qiymət verməsi onun irsinin və şəxsiyyətinin əhəmiyyətini bir daha təsdiqləyir. Xüsusilə vurğulamaq lazımdır ki, “A.S. Puşkin Mətnlər Korpusu” tədqiqata müqayisəli təhlil məqsədi ilə cəlb olunmuşdur. “A.S. Puşkin Mətnlər Korpusu”nun öyrənilməsi tədqiqatçıların işini xeyli asanlaşdırır və müxtəlif növ tədqiqatların imkanlarını nümayiş etdirir və yaratdığımız mətn korpusları üçün bir nümunə rolunu oynayır.

Korpus və linqvostatistik tədqiqatlara əhəmiyyətli töhfələr vermiş alimlər sırasında V.P. Zaxarov, A.A. Polikarpov, O.V.

Kukuşkina, A.A. Axundov, C.A. Rəhmanov, M.Ə. Mahmudov, K. Vəliyeva, R. Mahmudova, A.İ.İzotov, A.V. Zubov, M.A. Kajarova, E.A. Kazak, A.V. Kaşeyeva, M.V. Kopotyev və bir çox başqalarının adlarını qeyd etmək olar.

A.A. Polikarpovun ideyaları, Moskva Dövlət Universitetinin “Ümumi və Kompüter Leksikologiyası və Leksikoqrafiyası Laboratoriyası”nda (ÜKLLL) korpus menecerlərinin hazırlanması və mətn korpuslarının yaradılmasında böyük rol oynamışdır. Onun və O.V. Kukuşkinanın əməyi sayəsində A.S. Puşkin, A.N. Çexov və digər yazıçıların korpusları tərtib olunmuşdur.

A.A. Axundovun araşdırmaları Azərbaycanda linqvostatistikanın öyrənilməsinə mühüm töhfə vermişdir. Onun “Azərbaycan dilinin saıtlərinin statistik təhlili təcrübəsindən” adlı məqaləsi statistik və kompüter metodlarının tətbiqi ilə yazılmış ilk tədqiqat işlərindən biri hesab olunur. Sonralar A.A. Axundov bu sahədə araşdırmalarını genişləndirmiş və “Riyazi dilçilik” əsərini yazmışdır. Azərbaycanda linqvostatistik araşdırmaların inkişafı M.Ə. Mahmudov, K. Vəliyeva, M. Məmmədova və başqalarının adları ilə də bağlıdır. Belə ki, M.Ə. Mahmudov ölkədə ilk olaraq avtomatik dil təhlili, maşın tərcüməsi və statistik leksikoqrafiya sahəsində tədqiqatlar aparmışdır. O, Azərbaycan dili materialında mətnlərin formal təhlili problemini sistemli şəkildə öyrənmiş, dilin müxtəlif səviyyələrində formal işarələrin avtomatik işlənməsi üçün alqoritmlər qurmuş və onları hesablama maşınlarında tətbiq etmişdir. Azərbaycan dilçiliyində korpus araşdırmalarının əsasları isə C.A. Rəhmanov, M.A. Mahmudov, K. Vəliyeva və başqalarının işlərində əksini tapmışdır. Azərbaycanda korpus dilçiliyi üzrə ilk araşdırmalar C.A. Rəhmanova məxsusdur. Onun işində maşın tərcüməsi problemləri araşdırılmışdır.

Beləliklə, azərbaycanlı tədqiqatçıların əsərləri Azərbaycanda korpus dilçiliyi və linqvostatistika sahələrinin öyrənilməsinə xeyli töhfə vermişdir. Bu da mətn korpusunun yaradılması məsələsinin qoyulmasında mühüm rol oynamış və nəticədə hazırda Azərbaycanda korpus dilçiliyi sahəsində Azərbaycan dili elektron lüğət korpusları yaradılmışdır. Yuxarıda adı çəkilən alim və tədqiqatçıların elmi işləri öncəliklə yeni elmi istiqamətlərin yaranması ilə bağlı vacib məsələləri

qaldırması baxımından əhəmiyyətlidir. Lakin bu işlərdə ayrı-ayrı müəlliflərin mətn korpuslarının və azərbaycanlı müəlliflərin dil lüğətlərinin yaradılmasının əhəmiyyəti nəzərdən keçirilmir. . Həmçinin bu araşdırmalar əsasən nəzəri istiqamət daşıyır, dissertasiya işi isə Azərbaycan dili mətnlərinin ilk korpuslarının yaradılması, tərtib edilmiş konkordansların təhlili, müəllif dilinin fərqli xüsusiyyətlərinin araşdırılmasına həsr olunmuşdur.

**Tədqiqatın obyektı və predmeti.** Tədqiqatın obyektı linqvostatistik və korpus tədqiqi baxımından bədii mətnlərdir. Tədqiqatın predmeti isə tərtib olunmuş verilənlər bazaları, konkordanslar və mətn korpusları, həmin korpuslara və verilənlər bazasına (VB) daxil olan leksik vahidlər və onların işlənmə xüsusiyyətləridir. Əsasən tezlik xarakteri daşıyan bu konkordanslar kompüter leksikoqrafiyasının məhsulu hesab olunur.

**Tədqiqatın məqsədi və vəzifələri.** Dissertasiyanın məqsədi Azərbaycan müəlliflərinin (şair və yazıçıların) mətn korpuslarını, həmçinin tədqiq olunan bədii mətnlərə daxil olan vahidlərin konkordanslarını, həmin vahidlərin tezlik siyahılarını (lüğətlərini) tərtib etməkdir.

Bu məqsədə nail olmaq üçün aşağıdakı vəzifələri həll etmək qarşıya qoyulmuşdur:

- “Mətn” anlayışının müxtəlif şərhlərini təsvir və təhlil etmək;
- Mətn kateqoriyaları və komponentlərini təsvir etmək;
- “Mətn” və “diskurs” anlayışları arasındakı fərqi təhlil və müəyyən etmək;
- “Mətn” və “bədii mətn” anlayışları arasındakı fərqi aşkar etmək;
- Yeni fənlərarası sahələrin – kompüter dilçiliyi, linqvostatistika, korpus dilçiliyinin mənşəyini və üstünlüklərini təsvir etmək;
- Korpus dilçiliyi və linqvostatistika metodlarının tətbiq dairəsini müəyyənləşdirmək;
- A.P. Çexov və A.S. Puşkinin mətnlər korpuslarını təsvir etmək və bir nümunə kimi araşdırmaq;

- Tədqiqata “A.S. Puşkin Mətn Korpusu”nu cəlb etməklə mətn korpuslarının müxtəlif istiqamətli və genişmiqyaslı araşdırmalar aparmaq imkanlarını nümayiş etdirmək;

- A.S. Puşkinin “Kapitan qızı” povestini və onun Azərbaycan dilinə tərcüməsini təhlil etmək; orijinal və tərcümə dillərində situativ modallığın leksik vahidlərinin işlənməsini müqayisə etməklə fərqləri üzə çıxarmaq;

- Azərbaycan şair və yazıçılarının mətn korpuslarını yaratmaq;
- Tezlik lüğətləri və konkordanslar tərtib etmək;
- Token və söz formalarının sayını müəyyənləşdirmək;
- Ayrı-ayrı müəlliflərin dil xüsusiyyətlərini üzə çıxartmaq;
- Müəyyən bir müəllifin əsərlərində aşağı tezliyi və yüksək tezliyi olan leksikanın rolunu müəyyən etməkdir.

**Tədqiqat üsulları.** Dissertasiyada təsviri, komponent təhlili, korpus dilçiliyi, leksikoqrafik təhlil metodlarından, mətn emalının statistik və digər üsullarından istifadə olunmuşdur.

**Müdafiəyə çıxarılan əsas müddəalar:**

- Müəllif və bədii mətnlər korpuslarının təhlili nəticəsində bədii mətnin spesifik struktur xüsusiyyətləri müəyyən edilmişdir. Bu xüsusiyyətlər sintaktik mürəkkəbliyə, üslublararası vahidlərin mövcudluğu və üslubi baxımdan işarələnmiş leksikanın yüksək sıxlığı ilə ifadə olunur ki, bu isə onun mətnlərin tipologiyasında çoxsəviyyəli dil sistemi kimi statusunu dəqiqləşdirir;

- Müxtəlif janr mətnlərinin korpus təhlili əsasında bədii mətnin fərqləndirici xüsusiyyətləri aşkar edilmişdir. Bunlara ekspressiv leksikanın üstünlük təşkil etməsi, trop və üslubi fiqurlardan geniş istifadə, həmçinin danışmaq və publisistik leksika elementlərinin daxil edilməsi tendensiyası aid edilir ki, bu isə onun funksional hibridliyini təsdiq edir;

- İlk dəfə rus və Azərbaycan dillərində bədii mətnlərin korpus təhlili materialında mətn və diskurs arasındakı fərq empirik şəkildə təsdiqlənmişdir. Mətn nitq fəaliyyətinin nəticəsi kimi, diskurs isə kommunikativ situasiyada proses kimi qeydə

alınmışdır; yalnız mətnin statistik və tezlik baxımından emala tabe olduğu əsaslandırılmışdır;

- Rus dilinin xarici dil kimi tədrisində korpus məlumatları əsasında Azərbaycan dilinə tərcümələrindən ibarət tezlik lüğətləri və siyahılar (isim, fel, sifət və s.) hazırlanaraq tədris prosesinə tətbiq olunmuşdur. Bu, azərbaycanlı tələbələrin leksikanı mənimsəməsində effektivliyini artırmışdır;

- A.S. Puşkin, M.Ə. Sabir, M. Füzulinin mətnləri korpuslarının təhlili əsasında bədii mətnlərdə müxtəlif nitq hissələrinin və bəzi sintaktik konstruksiyaların istifadə qanunauyğunluqları müəyyən edilmişdir. Bu isə müəlliflərin üslub xüsusiyyətlərini formallaşdırmağa və mətnlərin daha dəqiq avtomatik işarələnmə və emalına şərait yaradır. Beləliklə, korpuslar vahidlərin işlənmə mənzərəsini tam şəkildə göstərir və onların istifadə dərəcəsini müəyyən etməyə imkan verir;

- İlk dəfə bədii mətn korpuslarının yaradılmasının və onların Azərbaycan Milli Korpusuna bədii ədəbiyyat və şifahi xalq yaradıcılığı əsasında formalaşdırılmış alt korpuslar kimi daxil edilməsinin vacibliyi əsaslandırılmışdır. Bu alt korpuslar presedent fenomenləri, mədəniyyətə xas reallıqları və müəllif üslub modellərini ehtiva edir ki, bu da korpusun analitik potensialını genişləndirir və onun fənlərarası tətbiqini mümkün edir;

- A.S. Puşkinin mətnlər korpusu müəllif korpusun nümunəsi kimi müqayisəli məqsədlərlə istifadə olunmuş və korpus dilçiliyi metodlarının araşdırmalarda yüksək effektivliyini təsdiqləmişdir. Puşkinin korpusuna tətbiq olunan metodlar analitik işin asanlaşdırılmasını, mətnin tezlik, üslubi və struktur xüsusiyyətlərinin aşkarlanmasını göstərir;

- Yaradılmış müəllif mətn korpusları müəyyən müəlliflərin dil xüsusiyyətlərini, o cümlədən leksiko-semantik üstünlüklərini və fərdi üslubi üsullarını aşkar etməyə imkan verir ki, bu da onları linqvopersonologiya təhlilinin aləti kimi əhəmiyyətli edir. Bir müəllifin mətnlərində dil bütöv bir sistem kimi tam şəkildə reallaşır. Belə korpuslar milli korpus

məktəbinin formalaşdırılması üçün xüsusi dəyər kəsb edir, xüsusilə Azərbaycan korpus dilçiliyinin formalaşma mərhələsində dil fərdiliyinin sistemli təsviri həm elmi, həm də tətbiqi əhəmiyyət daşıyır;

- Yaradılmış müəllif korpusları əsasında ayrıca müəllif dillərinin tezlik lüğətləri (M.Ə. Sabirin, M. Füzulinin, Nizami Gəncəvinin) tərtib edilmişdir ki, burada xarakterik mövzu sahələri, üstünlük təşkil edən nitq hissələri əks olunmuşdur. Bu işə leksik-üslubi təhlilin daha dərin səviyyədə aparılmasına şərait yaradır;

- Klassik poeziya və nəsr əsərlərinin orijinal və tərcümələrinin tezlik siyahıları və konkordanslarının təhlili nəticəsində (N. Gəncəvinin “Leyli və Məcnun” poemasının rus və Azərbaycan dillərinə tərcüməsi və Puşkinin “Kapitan qızı” əsərlərinin orijinal və tərcümə dillərində) mətnlərin leksik tərkibində mühüm fərqlər müəyyən edilmişdir. Tərcümə prosesində semantik tamlığına və üslubi çalarına, eləcə də müəyyən dil vahidlərinin tezliyinə təsir göstərən əvəzlənmələrin baş verdiyi aşkar olunmuşdur ki, bu da mətnin adaptasiya xüsusiyyətlərini üzə çıxarır və tərcümə üslubunun xarakterik cəhətlərini müəyyən etməyə imkan verir;

- A.S. Puşkinin, M.Ə. Sabirin və M. Füzulinin müəllif korpuslarının müasir rus və Azərbaycan ədəbi dili materialı ilə müqayisəsi əsasında onların dilinin fərqləndirici xüsusiyyətləri müəyyən edilmişdir. Müəlliflərin leksik-qrammatik xüsusiyyətləri və tezlik fərqləri aşkara çıxarılmışdır ki, bunlar həm tarixi-mədəni konteksti, həm də fərdi üslubu əks etdirir;

- İlk dəfə Azərbaycan dilində bədii mətn korpusları yaradılmış və təsvir edilmişdir: M.Ə. Sabirin mətnlər korpusu, Nizami Gəncəvinin “Xəmsə”sinin Azərbaycan dilinə tərcümələri korpusu, habelə M. Füzulinin poetik korpusu. Bu korpuslar əsasında tezlik lüğətləri və konkordanslar hazırlanmışdır ki, bu da Azərbaycan dilinin korpus dilçiliyinin inkişafına mühüm töhfə vermiş, leksikoqrafik və linqvostilistik

tədqiqat imkanlarını genişləndirmiş, həmçinin dil, tərcümə və digər sahələrdə yeni elmi araşdırmalar üçün əsas yaratmışdır;

- Müəllif korpuslarının korpus və linqvostatistik təhlili nəticəsində aşağı tezliyi olan leksikanın fərdi müəllif üslubunun göstəricisi kimi təsbit edilmişdir. Korpusda minimal tezliyi olan vahidlər, adətən, müəllifin fərdi leksikasını təşkil edir və unikal nominasiya, neologizm və ya idiolektizmlərdən ibarət olur.

**Tədqiqatın elmi yeniliyi.** Dissertasiyanın elmi yeniliyi, ilk növbədə, bədii mətnin xüsusi statusunun, onun korpus və linqvostatistik tədqiqatlarda əhəmiyyətinin müəyyən edilməsinə cəhd göstərilməsindədir. Bundan əlavə, elmi yenilik korpus dilçiliyi və linqvostatistik metodlarının ayrıca bir araşdırması çərçivəsində tətbiqi ilə yanaşı, Azərbaycan və qismən rus dillərində təqdim olunmuş bədii mətnlərin tədqiqində korpus linqvistikası və linqvostatistika metodlarının təhlilində istifadə olunmuşdur. Bu metodların vasitəsi (bu yanaşma) ilə müəyyən (konkret müəlliflərin) bir müəllifin dil xüsusiyyətlərini, bədii mətnin orijinal və tərcümə dilindəki spesifikasını, müəllif dilinin xarakterik cəhətlərini üzə çıxarmağa imkan verir. Elmi yeniliyin əsas əhəmiyyəti isə mətn korpuslarının yaradılması ilə əsaslandırılır.

Xüsusilə qeyd etmək lazımdır ki, Azərbaycanda korpus dilçiliyi və linqvostatistika sahəsində tədqiqatlar zəif və az sayda təqdim olunduğundan bu tip tədqiqatlar innovativ xarakter daşıyır. Daha dəqiq desək, Azərbaycan dili mətn korpusları ilə təmsil olunmamışdır. Bu tədqiqat Azərbaycan müəlliflərinin (yazıçı və şairlərinin), Azərbaycan dilinin mətnlər korpusunun formalaşmasına xidmət edəcək və Azərbaycan Milli Korpusunun yaradılmasına töhfə verəcəkdir.

Milli korpusda bədii ədəbiyyat nümunələri də yer alır, çünki məhz bədii ədəbiyyat əsərləri və şifahi xalq yaradıcılığı nümunələri mədəni əhəmiyyət kəsb edir və onlar precedent fenomenləri özündə cəmləşdirir.

Tezlik lüğətlərinin tərtibi xüsusi əhəmiyyət kəsb edir, çünki bu cür lüğətlərin tətbiqi maşın tərcüməsi sistemlərində məlumatların avtomatik emalı, referatlaşdırma, poliqrafiya, kriminalistika, habelə

anonim qeydlərin və mətnlərin müəllifliyinin müəyyənləşdirilməsi kimi sahələrdə mühüm rol oynayır.

Belə ki, Rusiyada Puşkinin və Çexovun mətnlərinin korpusları yaradılmış, İngiltərədə isə Şekspirin mətnlər korpusu formalaşdırılmışdır. Müəllif korpusları son dövrlərdə xüsusilə intensiv şəkildə yaradılır. Bu bərdə V.P. Zaxarov da öz elmi işində yazmışdır. Müəllif korpusları ayrı-ayrı fərdlərin dil xüsusiyyətlərini araşdırmağa imkan verir. Belə ki, bir fərdin mətnlərində dil vahid bir sistem kimi reallaşır. Müəllif korpuslarından əlavə, tədqiqatçı qarşıya qoyulan məqsəd və vəzifələrdən asılı olaraq, müxtəlif sorğulara cavab verən yeni korpuslar da tərtib edə bilər; yəni tədqiqatçı özü müəyyən korpusun tərtibçisi rolunu oynaya bilər. Məsələn, elmi mətnləri emal edən korpus yaradılmışdır. Beləliklə, təqdim olunan iş Azərbaycan elmi sahəsində müəllif korpuslarının yaradılmasında ilk tədqiqat hesab olunur.

**Tədqiqatın nəzəri və praktiki əhəmiyyəti.** Dissertasiyanın nəzəri əhəmiyyəti bir sıra anlayışları işıqlandırmaq, mətn haqqında elmi fikirləri genişləndirmək və bədii mətnin xüsusi statusunu müəyyən etməkdir. Bundan başqa, nəzəri əhəmiyyət korpus dilçiliyinin inkişafını, korpus və linqvostatistik metodlardan istifadənin üstünlüklərini əsaslandırmaqdan, ayrı-ayrı yazıçıların mətn korpuslarının yaradılmasının əhəmiyyətini müəyyənləşdirməkdən ibarətdir. Dissertasiya nəzəri və praktik leksikoqrafiya, tezlik lüğətləri və korpuslar sahəsinə də müəyyən töhfə verir.

İşin praktiki əhəmiyyəti ondan ibarətdir ki, əldə edilmiş nəticələr leksikologiya, mətn dilçiliyi və s. kurslarda, praktik məşğələlərdə istifadə oluna bilər. Bundan əlavə, yaradılmış tezlik lüğətləri və konkordanslar tədris və leksikoqrafik məqsədlər üçün istifadə edilə bilər. Yaradılmış mətn korpusları tədris və leksikoqrafik məqsədlər üçün istifadə edilə bilər. Bundan əlavə, bu tədqiqat Azərbaycanda korpus dilçiliyinin və linqvistik statistikanın inkişafına, mətn korpusunun yaradılmasına öz töhfəsini verəcəkdir.

Tədqiqatın praktiki əhəmiyyəti ondan ibarətdir ki, əldə olunmuş nəticələr leksikologiya, mətn dilçiliyi və digər kurslarda, praktik məşğələlərdə istifadə oluna bilər. Bundan əlavə, yaradılmış

tezlik lüğətləri və konkordanslar tədris və leksikoqrafik məqsədlər üçün tətbiq edilə bilər. Yaradılmış mətn korpusları da tədris və leksikoqrafik məqsədlər üçün istifadə oluna bilər. Bundan başqa, bu tədqiqat Azərbaycanda korpus dilçiliyi və linqvostatistikanın inkişafına, mətn korpuslarının yaradılmasına töhfə verəcəkdir.

**Aprobasiyası və tətbiqi.** Tədqiqatın əsas müddəaları iddiaçının bir sıra respublika və beynəlxalq elmi konfrans və simpoziumlardakı məruzələrində əks olunmuşdur. Bu dissertasiyanın mövzusu ilə bağlı 20 məqalə, tezis və konfrans materialları dərc edilmişdir.

**Dissertasiya işinin tamamlandığı müəssisənin adı.** Dissertasiya işi Azərbaycan Milli Elmlər Akademiyası Nəsimi adına Dilçilik İnstitutunun Süni intellekt və kompüter dilçiliyi şöbəsində aparılmışdır.

**Dissertasiyanın strukturu və həcmi və işarə ilə ümumi həcmi.** Dissertasiya giriş, mündəricat, üç fəsil, nəticə, istifadə olunan ədəbiyyat siyahısı, əlavələr və ixtisarlardan ibarətdir. Giriş – 12 səhifə, 20 618 işarə, I Fəsil – 44 səhifə, 73 797 işarə, II Fəsil – 32 səhifə, 56 402 işarə, III Fəsil – 62 səhifə, 96 007 işarə, Nəticə – 7 səhifə, 12 528 işarə. Dissertasiyadakı simvolların ümumi sayı 259 352 işarə təşkil edir.

## TƏDQIQATIN ƏSAS MƏZMUNU

Dissertasiyanın **Giriş** hissəsində mövzunun aktuallığı və işlənmə dərəcəsi əsaslandırılır, tədqiqatın predmeti və obyekt, məqsəd və vəzifələri, tədqiqat metodları, müdafiyyə çıxarılan əsas müddəalar müəyyənləşdirilmiş, tədqiqatın elmi yeniliyi, nəzəri və praktiki əhəmiyyəti, tədqiqat işinin aprobasiyası və tətbiqi, dissertasiya işinin yerinə yetirildiyi təşkilatın adı, dissertasiyanın struktur bölmələrinin ayrılıqda həcmi və işarə ilə ümumi həcmi haqqında məlumat təqdim edilmişdir.

Birinci fəsil **“Mətn” və “bədi mətn” anlayışları. Bədi mətnin xüsusi statusu” adlanır** və 5 yarım fəsildən ibarətdir. “Filologiyanın obyekt kimi mətn. **“Mətn” problematikasının araşdırılmasında müxtəlif yanaşmalar”** adlı yarım fəsində “mətn” anlayışının

Azərbaycan, rus və xarici dilçilikdə yaranması və tədqiq olunması nəzərdən keçirilir.

Mətn dilçilikdə ən mühüm və əsas terminlərdən biridir. Eyni zamanda, mətn filologiyasının başlıca tədqiqat obyektidir. Yu.M. Lotmanın qeyd etdiyi kimi, mətn “*şübhəsiz ki, humanitar elmlərdə ən çox istifadə olunan terminlərdən biridir*”.<sup>1</sup> “*Mətn (latınca *textus* - parça, əlaqə) əsas xassələri ardıcılıq və bütövlük olan semantik əlaqə ilə birləşən simvolik vahidlərin ardıcılığıdır*”.<sup>2</sup> Bu tərifdən belə anlaşılır ki, mətndəki simvolik vahidlər ümumi semantik komponentə malik olmalıdır, bu komponentin əsasında onlar birləşir və nəticədə mətn formalaşır. Aydınadır ki, hər işarə toplusu mətn hesab oluna bilməz. Mətn olmaq üçün iki əsas şərt – bağlılıq (uyğunluq) və bütövlük tələb olunur. Həmçinin qeyd edilməlidir ki, “mətn” sözünün etimologiyası da onun əsas prinsiplərindən birini – birləşməni, yəni semantik uyğunluğunu əks etdirir.

Mətn mürəkkəb və çoxşaxəli sistemdir, buna görə də elm aləmində onun şərhinə dair müxtəlif yanaşmalar mövcuddur.

Məsələn, L.Q. Babenko qeyd edir ki, “*mətn bütün dil vahidlərinin reallaşdığı şifahi nitq əsəridir*”.<sup>3</sup>

Z.Y. Turayevanın fikrincə, mətn “*müxtəlif leksik, məntiqi və qrammatik əlaqə növləri ilə birləşmiş, mütəşəkkil və yönləndirilmiş məlumatı müəyyən şəkildə ötürməyə qadir olan nizamlı cümlələr toplusudur. Mətn mürəkkəb bütövdür, struktur-semantik vəhdət kimi fəaliyyət göstərir*”.<sup>4</sup> Tədqiqatçı mətni mürəkkəb sistem hesab edir və mətnin bütün vahidlərinin struktur-semantik birliyini vurğulayır.

---

<sup>1</sup> Лотман, Ю.М. Семиосфера / Ю.М. Лотман. – СПб.: Искусство-СПб, – 2000. – 704 с.

<sup>2</sup> Лингвистический энциклопедический словарь: [Электронный ресурс] / URL: <http://tapemark.narod.ru/lcs/index.html>

<sup>3</sup> Бабенко, Л.Г. Лингвистический анализ художественного текста [Текст] / Л.Г. Бабенко, И.Е. Васильев, Ю.В. Казаров. – Екатеринбург: Изд-во Урал, унта, 2000. – 534 с.

<sup>4</sup> Тураева, З.Я. Лингвистика текста (Текст: Структура и семантика): [Электронный ресурс] / URL: [https://www.phantastike.com/linguistics/lingvist\\_teksta/djvu/view/](https://www.phantastike.com/linguistics/lingvist_teksta/djvu/view/)

Q. Kazımov mətnin qurulmasında tema-rematik münasibətlərin əhəmiyyətini xüsusi vurğulayır. Onun fikrincə, cümlənin kommunikativ əsasını təşkil edən yeni informasiya düşüncənin dərinləşməsinə və inkişafına gətirib çıxarır, nəticədə isə ayrı-ayrı ifadələr bir-biri ilə birləşərək mətnin tematik inkişafını, yəni mətnin yaradılmasını şərtləndirir. Alim öz fikrini belə ümumiləşdirir: mətn – ümumi fikir əsasında birləşmiş struktur və mənaca bütövlük təşkil edən cümlələrin vəhdətidir.<sup>5</sup>

T.M. Dridze isə hesab edir ki, mətn – mürəkkəb bir işarə təşkil edən bütöv kommunikativ vahiddir. Bu sistemin daxilindəki elementlər isə vahid struktura malik olmaqla müəyyən iyerarxiyaya tabe edilir.<sup>6</sup>

N.B. Məmmədli isə qeyd edir ki, “*mətn strukturun, kompozisiyanın, qrammatik-kommunikativ tamlığın və məzmunun birliyini ifadə edir*”<sup>7</sup>. Digər təriflərə görə də mətn «çoxplanlı sistem» və struktur-semantik bütövlüyə malik mürəkkəb bir vahiddir.

N.F. Alefirenko isə qeyd edir ki, mətn “*hər hansı dildə yazılan və deyilənlərin məcmusudur*”.<sup>8</sup> Deməli, mətn müxtəlif dil vahidlərindən təşkil olunur və çoxsəviyyəli mürəkkəb sistemdir.

Beləliklə, daha dəqiq təriflər Z.Y. Turayevaya, Q. Kazımova və N.B. Məmmədliyə məxsusdur. N.F. Alefirenko da əhəmiyyətli mövqe bildirir: o, mətni yazılı və şifahi formalara ayırır.

Bu yarımfəsildə mətn araşdırmasına dair müxtəlif yanaşmalar nəzərdən keçirilmiş və müəyyən olunmuşdur ki, əsas aspekt məhz linqvistik, yaxud linqvosentrikdir.

**“Mətnin komponentləri və kateqoriyaları”** adlı yarımfəsildə mətnin tərkib hissələri, elementləri və kateqoriyaları araşdırılır və

---

<sup>5</sup> Kazımov, Q.Ş. Müasir Azərbaycan dili. Sintaksis / Q.S. Kazımov. – Bakı: Ünsiyyət, – 2000. – 496 s.

<sup>6</sup> Дридзе, Т.М. Текстовая деятельность в структуре социальной коммуникации: проблемы семиосоциопсихологии [Текст] / Т.М. Дридзе. – М.: Наука, – 1984. – 268 с.

<sup>7</sup> Mətn dilçiliyi problemləri / N.B. Məmmədli. – Bakı: Zərdabi Nəşr MMC, – 2023. – 260 s.

<sup>8</sup> Alefirenko, H.Ф.: [Электронный ресурс] / URL: <https://core.ac.uk/download/pdf/151213879.pdf>

cümlənin əsas mətn vahidi olduğu müəyyən edilir. Yanaşmadan asılı olaraq, tədqiqatçılar cümləni müxtəlif cür adlandırırlar: “*minimal kommunikativ vahid*” (Qroşenkova)<sup>9</sup>, “*fikir zəncirində kiçik predikativ halqa*” (N.B. Məmmədli)<sup>10</sup> və s. Mətnin əsas kateqoriyaları isə bağlılıq və bütövlükdür.

“*Mətn və diskurs*” adlı yarım-fəsildə “mətn” və “diskurs” anlayışları məsələsinə də toxunulur, onların fərqləndirilməsi tədqiqatın obyektini müəyyənləşdirilir. Bəzi tədqiqatçılar bu iki anlayışı sinonim hesab edir, digərləri isə fərqləndirirlər. Alt fəsildə “diskurs” anlayışının yaranma mənbələri təsvir olunmuş, müxtəlif tədqiqatçıların baxışları nəzərdən keçirilmişdir. Belə ki, diskurs – “*ekstralingvistik (praqmatik, sosial-mədəni, psixoloji və digər) amillərlə sıx əlaqədə olan mətn; hadisəvi aspektdə nəzərdən keçirilən mətn; məqsədyönlü sosial fəaliyyət kimi nəzərdən keçirilən nitq, insanların qarşılıqlı əlaqəsində və şüur (kognitiv proseslər) mexanizmlərində iştirak edən komponent*” kimi müəyyənləşdirilir.<sup>11</sup> *Diskurs sadəcə mətn deyil, “həyatla bağlı olan nitqdır”*<sup>12</sup>, yəni o, nitq və situasiyanın mürəkkəb və vəhdət vahidi kimi çıxış edir.

“*Bədii mətn. Bədii mətnin lingvistik təbiəti*” adlı yarım-fəsildə bədii mətnə dair müxtəlif baxış və şərhlər, “mətn” və “bədii mətn” anlayışları arasındakı fərq və oxşarlıqlar da araşdırılmaqdadır.

Bədii mətn – müəyyən dildə yaradılmış, bədii nitqin orijinal əsəri olaraq mürəkkəb və bütöv sistemdir. O, həm xalqın, həm də müəllifin dilinin xüsusiyyətlərini əks etdirir və ədəbi-bədii üslubdan istifadəsi səbəbindən xüsusi statusa malikdir. Bu üslub funksional

---

<sup>9</sup> Groşenkova, В.А. Обучение чтению в специальной (коррекционной) школе: [Электронный ресурс] / URL: <https://citoweb.yyspu.org/link1/metod/met143/node3.html>

<sup>10</sup> Mətn dilçiliyi problemləri / N.B. Məmmədli. – Bakı: Zərdabi Nəşr MMC, – 2023. – 260 s.

<sup>11</sup> Клокова, А.Г. Понятие дискурса в отечественном и зарубежном языкознании: [Электронный ресурс] / URL: <https://elib.bsu.by/handle/123456789/199391>

<sup>12</sup> Лингвистический энциклопедический словарь: [Электронный ресурс] / URL: <http://tapemark.narod.ru/les/index.html>

üslubların ən zəngini olmaqla digər bütün üslubların elementlərini əhatə edir.

Bu xüsusiyyət bədii mətni lingvostatistik, korpus və digər tədqiqat növləri üçün dəyərli obyektə çevirir. Deməli, “mətn” ümumi anlayış, “bədii mətn” isə onun növlərindən biridir.

I fəslin **“Bədii mətnin janrları”** adlı yarımfəslində isə ədəbi növlər və janrlar nəzərdən keçirilir. Tədqiqatın obyektı bədii mətnlər, predmeti isə onların tərkib vahidləri, konkordanslar və korpuslar olduğundan, janrların nəzərdən keçirilməsi onların struktur bütövlüyünü və dil vahidlərinin işlənmə xüsusiyyətlərini anlamaq baxımından olduqca vacibdir.

**“Korpus dilçiliyi və lingvostatistika: bədii mətnlərin və dilin mətn korpusları vasitəsilə tədqiqi”** adlı II fəsil 5 yarımfəsildən ibarətdir. **“Kompüter dilçiliyi. Elmin formalaşması”** adlı yarımfəsildə yeni fənlərəarası elm – kompüter dilçiliyi nəzərdən keçirilir. Texnologiyanın inkişafı ilə elmin yeni sahələri meydana gəlmişdir ki, bunlardan biri də kompüter dilçiliyidir.

“Kompüter dilçiliyi” anlayışı geniş mənada şərh olunur. Kompüter dilçiliyi digər elmi sahələrin kəsişməsində yaranmış yeni bir istiqamət kimi qiymətləndirilə bilər. Bundan başqa, dilçilikdə kompüterlərin tətbiqi və istifadəsi ilə bağlı bütün məsələlər (istər praktiki, istərsə də nəzəri) kompüter dilçiliyinin tədqiqat obyektı hesab olunur.

Kompüter dilçiliyinin yaranması və inkişafı bilavasitə elektron hesablama maşınlarının meydana çıxması ilə bağlıdır. XX-ci əsrin 50-ci illərində elmi təfəkkürdə bir növ “inqilab” baş vermiş, elmi düşüncədə çevriliş olmuşdur. Elektron-hesablama maşınlarının yaranması elmin inkişafını müəyyən etmiş, yeni termin və sahələr yaratmış, müxtəlif sahələrin kəsişməsində olan tədqiqatların daha dəqiq nəticələr verdiyini aşkar etmiş və bunlar kompüter dilçiliyinin yaranmasına təkan vermişdir.

Qeyd etmək lazımdır ki, elmi aləmdə bu sahənin son adı ilə bağlı yekdil fikir yoxdur. Bu elm sahəsi “riyazi dilçilik”, “struktur

dilçilik”, “kibernetik dilçilik”, “mühəndislik dilçiliyi”, “korpus dilçiliyi” və s. adlandırılmışdır.<sup>13</sup>

K. Boyarski isə qeyd edir ki, yaranmış yeni elm sahəsi ilkin mərhələdə müxtəlif cür adlandırılmış, lakin son nəticə olaraq “kompüter dilçiliyi” termini adı altında tanınmışdır.<sup>14</sup> K. Vəliyevanın sözlərinə görə, kompüter dilçiliyi təbii dilçiliyin bir bölməsi olub, kompüter proqramları, dil və təfəkkür vasitəsilə informatikanın dilçilik əsaslarını öyrənir, həmçinin dilin modelləşdirilməsini araşdıraraq ümumi dili formalaşdırır və təkmilləşdirir. Bununla yanaşı, o, süni intellektin əsas problemlərindən hesab olunan məsələlərin həllini də qarşıya məqsəd qoyur. Buraya, ilk növbədə, təbii dillərin interfeysi probleminin, eləcə də informasiya sistemlərinin modelləşdirilməsi məsələsinin həlli daxildir.<sup>15</sup>

Fikrimizcə, M.Ə. Mahmudov, K.K. Boyarski və K. Vəliyeva tərəfindən istifadə olunan “kompüter dilçiliyi” termini yeni sahənin müxtəlif aspektlərini əhatə edir və fərqli sahələrdən olan tədqiqatçıların tələblərinə cavab verir. Xüsusilə qeyd etmək lazımdır ki, bəzi sahələr müstəqil elm kimi inkişaf edə bilər. Məsələn, korpus dilçiliyi kompüter dilçiliyindən yaranmışdır.

**“Korpus dilçiliyi”** adlı yarım-fəsildə korpus dilçiliyinin yaranma tarixi, əsas anlayışları araşdırılır, onun terminoloji aparatı müəyyən və təsvir edilir, mətn təhlilində mühüm əhəmiyyət kəsb edən metodlar müzakirə edilir.

Korpus dilçiliyi kompüter dilçiliyi və linqvostatistikanın metodlarından istifadə etməklə mətn korpuslarının yaradılması və onların təhlili ilə məşğul olan dilçilik sahəsidir. Korpus dilçiliyinin terminoloji aparatına aşağıdakılar daxildir: mətn korpusu, konkordans, korpus meneceri və s.

---

<sup>13</sup> Боярский, К.К. Введение в компьютерную лингвистику / К.К. Боярский. Учебное пособие. – СПб: НИУ ИТМО, – 2013. – 72 с.

<sup>14</sup> Vəliyeva, K. Kompüter dilçiliyinə giriş / K. Vəliyeva. – Bakı: Elm və Təhsil, – 2018, – 384 s.

Korpus (lingvistik və ya dil korpusu) – müəyyən qaydalar əsasında müəyyənləşdirilmiş və tərtib olunmuş mətnlər toplusunun elektron-informasiya sistemidir: *“Linqvistik mətnlər korpusu termini xüsusi linqvistik problemləri həll etmək üçün nəzərdə tutulmuş böyük, elektron şəkildə təqdim edilmiş, vahid, strukturlaşdırılmış, etiketlənmiş, filoloji cəhətdən səlahiyyətli dil məlumatları massividir”*.<sup>16</sup>

Qeyd etmək lazımdır ki, korpus dilçiliyi metodları XIII əsrə qədər gedib çıxır: *“elektron dövründən əvvəlki korpuslar XII əsrdən 1960-cı illərə qədər formalaşmışdır”*.<sup>17</sup> Yəni, dilçilikdə “korpus” anlayışının istifadəsi yalnız elektron dönməindən əvvəlki son dövrlərə təsadüf edir. Bu dövrdən əvvəl isə «korpus» anlayışı dilçilik tərəfindən tanınmır və onun terminoloji aparatına daxil edilmirdi.

Korpusların əsas funksiyası – dilin öyrənilməsi və tədqiqi işlərini asanlaşdırmaq, qaydaları müəyyən etmək, müxtəlif tipli mətnlərin xüsusiyyətlərini kəmiyyət məlumatları əsasında araşdırmaq və korpus vahidlərinin istifadəsinin tam siyahısını təqdim etməkdir.

Mətn korpusu eləcə də korpus dilçiliyinin yaratmağa çalışdığı əsas obyektlərdən biridir. Əslində isə mətn korpusu “bir tərəfdən, onun obyektini kimi çıxış edir, yəni korpus dilçiliyi və digər dilçilik sahələri üçün ilkin nitq materialıdır”<sup>18</sup>; digər tərəfdən isə mətn korpusu korpus dilçiliyinin fəaliyyətinin nəticəsidir. Beləliklə, bu ikili xarakter korpus dilçiliyinin həm korpusların yaradılmasına, həm də tədqiqatçının qarşısına qoyulan müəyyən məsələlərin həllində mətn korpuslarından istifadəyə yönəldiyini göstərir.

Korpus dilçiliyi üçün “işarələmə” (“annotasiya”), “metaişarələmə”, “lemmatizasiya” və “tokenizasiya” kimi terminlər də mühüm əhəmiyyət kəsb edir. İşarələmə – mətnlər və söz işlənmələri

---

<sup>16</sup> Захаров, В.П. Корпусная лингвистика: Учебно-метод. пособие. – СПб., 2005. – 48 с.

<sup>17</sup> Солнышкина, М.И. История развития корпусной лингвистики (на примере англоязычных корпусов), URL: <https://cyberleninka.ru/article/n/istoriya-razvitiya-korpusnoy-lingvistiki-na-primere-angloyazychnyh-korpusov>

<sup>18</sup> Бойкова, С.Н. Введение в корпусную лингвистику: [Электронный ресурс] / URL: <https://core.ac.uk/download/pdf/75999967.pdf>

barədə əlavə linqvistik və metalingvistik məlumat deməkdir<sup>19</sup>. “*Metaişarələmə isə mətn haqqında mühüm ekstralingvistik informasiyanı (janr, müəllif, işlənmə sahəsi, mövzu və s.) təqdim edir*”.<sup>20</sup> Metaişarələmənin tərifindən də göründüyü kimi, o, mətnin janrı, müəllifi, mövzusu və digər xüsusiyyətləri haqqında əhəmiyyətli ekstralingvistik məlumat verir ki, bu da mətnin yalnız dil quruluşu ilə məhdudlaşmayan konteksti və özəlliklərini daha dərindən anlamağa imkan yaradır.

“*Lemmatizasiya – sözün digər sözformalarından çıxış edərək onun ilkin (başlanğıc) formasının yaradılması prosesidir*”<sup>21</sup>, tokenizasiya isə mətni ayrıca vahidlərə ayırmağı nəzərdə tutur. Bu proses olmadan avtomatlaşdırılmış emal mümkün deyil<sup>22</sup>. Yəni, tokenizasiya – mətnin tokenlərə və sözformalara bölünməsi deməkdir<sup>23</sup>.

Qeyd etmək lazımdır ki, Azərbaycanda korpus dilçiliyi üzrə, tədqiqatlar azdır. Azərbaycan dili mətn korpusları ilə təmsil olunmayıb, lakin elektron lüğətlər korpusu vardır. Korpusların olmaması dilin real mətnlər əsasında dərin təhlil imkanlarını məhdudlaşdırır, bu isə avtomatik mətn təhlili, məşin tərcüməsi və digər yüksək texnologiyaların inkişafını ləngidir. Azərbaycan dilində tam korpusların yaradılması dilçilik tədqiqatları üçün yeni imkanlar açacaq və ölkədə korpus dilçiliyinin inkişafını sürətləndirəcəkdir.

Xüsusi korpuslar arasında müəllif korpusları ayrıca qeyd olunmalıdır. Son dövrlərdə onların yaradılması daha intensiv xarakter almışdır, çünki bu, fərdi müəllif dilin xüsusiyyətlərini araşdırmağa

---

<sup>19</sup>Лозовая, А.А.: [Электронный ресурс] / URL: [http://elib.rshu.ru/files\\_books/pdf/rid\\_647a25806a7e42d5b628f117f8c5f6e1.pdf](http://elib.rshu.ru/files_books/pdf/rid_647a25806a7e42d5b628f117f8c5f6e1.pdf)

<sup>20</sup> Orada.

<sup>21</sup>Токенизация: [Электронный ресурс] / URL: [https://ozlib.com/1010207/literatura/osnovnye\\_protседury\\_obrabortki\\_estestvennog\\_o\\_yazyka\\_tokenizatsiya\\_lemmatizatsiya\\_stemming\\_parsing#google\\_vignette](https://ozlib.com/1010207/literatura/osnovnye_protседury_obrabortki_estestvennog_o_yazyka_tokenizatsiya_lemmatizatsiya_stemming_parsing#google_vignette)

<sup>22</sup> Orada.

<sup>23</sup>Гречачин, В.А. Понятие токенизации в корпусной лингвистике: [Электронный ресурс] / URL: [https://www.elibrary.ru/download/elibrary\\_25307213\\_89263284.pdf](https://www.elibrary.ru/download/elibrary_25307213_89263284.pdf)

imkan verir. Belə ki, bir şəxsin mətnlərində dil vahid sistem kimi reallaşır.

Korpus dilçiliyi metodlarının üstünlüklərini xüsusi vurğulamaq lazımdır. Bu metodlardan istifadə etməklə linqvistik, linqvokulturoloji, sosioloji və digər müxtəlif məsələlər həll olunur, fərqli müəlliflərin bədii əsərləri və onların tərcümələri müqayisəli təhlil edilir, müəllif üslubu xüsusiyyətləri müəyyənləşdirilir, vahidlərin tezliyi tədqiq olunur, həmçinin sözlərin, söz formalarının, qrammatik kateqoriyaların və söz birləşmələrinin morfoloji və kontekstual parametrləri nəzərə alınmaqla seçimi və axtarışı həyata keçirilir.

II fəslin **“Milli korpuslar”** adlı yarım fəslində “milli korpus” anlayışı, onun əhəmiyyəti və bir neçə ən çox representativ korpuslar da təsvir edilir. Ən vacib məqam isə Azərbaycan yazıçı və şairlərinin mətn korpuslarının və Azərbaycan dilinin Milli Korpusunun yaradılmasının aktuallığı məsələlərini qaldırılmasıdır. Azərbaycan dilinin Milli Korpusunun yaradılması strateji baxımdan olduqca mühüm məsələdir.

“M.V. Lomonosov adına Moskva Dövlət Universitetinin Filologiya fakültəsinin **“Ümumi və Kompüter Leksikologiyası və Leksikoqrafiyası Laboratoriyası” (ÜKLLL) haqqında məlumat**” adlı yarım fəsilə M.V. Lomonosov adına Moskva Dövlət Universitetinin Filologiya fakültəsinin “Ümumi və Kompüter Leksikologiyası və Leksikoqrafiyası Laboratoriyası” (ÜKLLL) təsvir edilir. Burada təbii dilin nəzəri tədqiqatları və öyrənilməsi aparılır. ÜKLLL-də aparılan araşdırmalar və korpus menecerləri haqqında da məlumat verilir. Tədqiqat materialları sırasına A.S. Puşkinin “Kapitan qızı” əsəri və “A.S. Puşkinin mətnlər korpusu” daxil edildiyindən, onların nəzərdən keçirilməsi əsaslandırılır.

Bundan əlavə, II Fəslin **“Linqvostatistika”** adlı yarım fəslində yeni sahə – linqvostatistika və ya kvantitativ dilçilik də nəzərdən keçirilmişdir.

Linqvostatistika (linqvistik statistika) – *“1) geniş mənada: dilçilikdə statistik metodların tətbiqi sahəsi; 2) dar mənada: əsasən*

*mətn daxilində dil vahidlərinin statistik paylanma tipləri ilə bağlı olan linqvistik materiala aid bəzi riyazi problemlərin öyrənilməsidir.*”<sup>24</sup>.

Linqvostatistik metodlar xüsusilə bədii mətnlərin tədqiqində məhsuldardır. Onlar dil vasitələrinin istifadəsinin xarakterini qiymətləndirməyə imkan verir. Bu metodlarla nitq müxtəlifliyi, söz və söz formalarının tezliyi müəyyən olunur, tokenlər hesablanır, müəllifin dil xüsusiyyətləri və bədii mətnin orijinal və tərcümə dilindəki spesifikliyi müəyyən edilir. Bu isə konkret müəlliflərin dil korpuslarının yaradılmasını asanlaşdırır.

***“Bədii mətnlərin təhlilində korpus dilçiliyi metod və texnologiyalarının tətbiq dairəsi. Azərbaycan müəlliflərinin (yazıçı və şairlərinin) bədii mətnlərinin korpuslarının yaradılması”*** adlı III fəsil 8 yarım fəsildən ibarətdir. ***“Bədii mətn materialı əsasında dilin öyrənilməsində korpus dilçiliyi və linqvostatistika metodlarının tətbiqi”*** adlı yarım fəsildə korpus dilçiliyi metod və texnologiyalarının tətbiq dairəsi müəyyən edilir.

Qeyd olunur ki, korpus dilçiliyi və linqvostatistika metod və texnologiyaları yalnız tədqiqatçılar tərəfindən deyil, həm də müəllim və tələbələr tərəfindən də yeni dil öyrənmək məqsədilə istifadə edilir.

Beləliklə, korpus dilçiliyi metod və texnologiyalarının tətbiqi sayəsində qısa müddətdə rus dilində ən çox tezliyi olan sözlərin Azərbaycan dilinə tərcüməsindən ibarət lüğət hazırlanmış, ən çox işlənən isim, fel, sifət və digər nitq hissələrinin siyahıları və onların Azərbaycan dilinə tərcümələri tərtib edilmişdir. Bu yanaşma və metod rus dilinin xarici dil (və ya ikinci dil) kimi öyrənilməsinin effektivliyini əhəmiyyətli dərəcədə artırır və dil biliklərinin səviyyəsini yüksəltməyə xidmət edir. Qeyd etmək lazımdır ki, kontekstlər Rus dilinin Milli Korpusundan götürülmüşdür və onlar bədii ədəbiyyat nümunələridir, yəni araşdırmada istifadə olunan materiallar korpusda təqdim edilmiş bədii mətnlərin kontekstlərindən ibarətdir. Bu isə tədqiqat baxımından mühüm əhəmiyyət kəsb edir.

---

<sup>24</sup> Лингвистический энциклопедический словарь [Электронный ресурс] / URL: <http://tapemark.narod.ru/les/index.html>

***“A.S. Puşkinin “Kapitan qızı” əsəri materialı əsasında linqvostatistik tədqiqat”*** adlı yarım fəsil həmin əsərin təhlilinə həsr edilir. Bu təhlil orijinal və Azərbaycan dilinə tərcümə arasındakı fərqlərin müəyyənləşdirilməsi məqsədilə aparılır. Qeyd etmək lazımdır ki, situativ modallıq subyektin hərəkətə münasibətini ifadə edir. Məsələn, “o bilər”, “o etməlidir”, “o istəyir”.

***“A.S. Puşkinin mətn korpusu üzrə “situativ modallıq” sahəsinin mikrosahələrini adlandıran bəzi isimlərin və əsas predikatların təhlilinin nəticələri”*** adlı yarım fəsil ÜKLLL-nin “A.S. Puşkin mətnlər korpusu” üzrə aparılan işlərin nəticələri təsvir edilir. III fəslin ***“Korpus məlumatları əsasında müasir dil və Puşkin dilində bəzi göstəricilərin işlənməsinin müqayisəli təhlili”*** adlı yarım fəslində isə Puşkinin dil xüsusiyyətlərinin müəyyən edilməsi və onun dilinin müasir ədəbi rus dilindən fərqləndirilməsi haqqındadır.

Xüsusulə qeyd etmək lazımdır ki, A.S. Puşkinin mətnlər korpusu dissertasiyada nümunə kimi istifadə olunur: onun materialı əsasında tədqiqat yaradılan korpusların güclü və zəif tərəflərini qiymətləndirməyə və onların tədqiqat potensialını nümayiş etdirməyə imkan verəcəkdir.

III fəslin ***“Azərbaycan yazıçı və şairlərinin mətn korpuslarının yaradılması”*** adlı yarım fəslində ayrı-ayrı müəlliflərin mətn korpuslarının yaradılmasının əhəmiyyəti qeyd olunur, həmçinin Azərbaycan müəllif korpuslarının yaradılmasının önəmi vurğulanır. Bu isə Azərbaycan dilinin Milli Korpusunun yaradılması baxımından vacibdir.

Qeyd etmək lazımdır ki, elmi aləmdə artıq müxtəlif müəlliflərin mətn korpusları mövcuddur. Məsələn, rus dilçiliyində A.S. Puşkinin mətnlər korpusu və A.P. Çexovun bədii mətnlər korpusu, İngiltərədə isə Şekspirin mətnlər korpusu və digərləri vardır. Araşdırma məqsədlərinə uyğun olaraq tədqiqatçı öz korpusunu da yarada bilər. Məsələn, elmi məqalələrin avtomatik təhlili üçün korpus yaradılmışdır.<sup>25</sup>

---

<sup>25</sup>Callegari, Elena Xhura, Desara A corpus for Automatic Article Analysis: [Электронный ресурс] / URL: <https://aclanthology.org/2022.clib-1.2.pdf>

Azərbaycan müəlliflərinin mətn korpuslarının yaradılması Azərbaycan dilinin Milli Korpusunun formalaşdırılması üçün də vacibdir. Belə ki, milli korpus dilin bütün stil və janrlarını, o cümlədən bədii ədəbiyyat və folkloru, mədəni əhəmiyyət daşıyan və precedent fenomenləri ehtiva edən mətnləri əhatə edir.

Həmçinin qeyd etmək lazımdır ki, tezlik lüğətlərinin tərtibi xüsusi dəyərə malikdir, çünki onların tətbiqi avtomatik mətn emalı, məşın tərcüməsi, referatlaşdırma, poliqrafiya, kriminalistika, anonim yazı və mətnlərin müəlliflərinin müəyyən edilməsi kimi sahələrdə böyük rol oynayır.

***“Nizami Gəncəvinin “Xəmsə” əsərlərinin mətnlər korpusunun yaradılması”*** adlı yarımfəsildə N. Gəncəvinin “Xəmsə”-sinin Azərbaycan dilinə tərcüməsi əsasında mətnlər korpusunun yaradılma prosesi təsvir edilir. Aparılmış korpus və linqvostatistik tədqiqatlar nəticəsində ümumi və hər poema üzrə konkordanslar, tezlik lüğətləri, eləcə də “Leyli və Məcnun” poemasının tərcüməsi əsasında etnospesifik leksikanın lüğəti tərtib edilmişdir. Hazırda həmin etnospesifik leksika lüğətinə 588 vahid daxil edilmişdir<sup>26</sup>. “Xəmsə” mətnlər korpusu üzrə kəmiyyət göstəriciləri müəyyənləşdirilmişdir: 290 476 token və 45 453 söz forması. Korpus üzrə ən çox işlənən vahidlər aşkar olunmuşdur: “bir” – 8193 işlənmə, “bu” – 4942 işlənmə, “o” – 4473 işlənmə, “hər” – 2617 işlənmə, “ki” – 2079 işlənmə; “nə” – 1830 işlənmə, də – 1802 işlənmə, “da” – 1489 işlənmə; “sən” – 1488 işlənmə, “mən” – 1341 işlənmə.

Bundan əlavə, “Leyli və Məcnun” poemasının Azərbaycan və rus dillərinə tərcümələrinin müqayisəli təhlili aparılmış, şəxs əvəzliliklərinin, predikatların, dini və etnospesifik leksikanın tezliyi qiymətləndirilmişdir. Ərəb mənşəli sözlərin funksionallığı müəyyən

---

<sup>26</sup>Халилова, Г.А. «Развитие азербайджанской корпусной лингвистики. Создание корпуса текстов азербайджанских авторов (поэтов и писателей) и национального корпуса азербайджанского языка». Вестник Дагестанского государственного университета. Серия2: Гуманитарные науки 2024 / 2. [Электронный ресурс] / URL: <https://vestnik.dgu.ru/pol.aspx?razdel=2&artId=4670>

edilmiş, “Allah”, “yaradan”, “tanrı” və s. sözlərin tezlik fərqləri aşkarnamışdır.

Korpus üzrə ən çox işlənən on vahid bunlardır: bir – 8193 işlənmə, “bu” – 4942 işlənmə, “o” – 4473 işlənmə, “hər” – 2617 işlənmə, “ki” – 2079 işlənmə; “nə” – 1830 işlənmə, “də” – 1802 işlənmə, “da” – 1489 işlənmə, “sən” – 1488 işlənmə, “mən” – 1341 işlənmə.

Bundan başqa, «Leyli və Məcnun» poemasının Azərbaycan və rus dillərinə tərcümələrinin müqayisəli təhlili aparılmış, şəxsi əvəzliliklərin, predikatların, dini və etnospesifik leksikanın işlənmə tezliyi araşdırılmışdır. Ərəb mənşəli sözlərin funksionallaşma xüsusiyyətləri müəyyən edilmiş, “Allah”, “yaradan”, “tanrı” və digər sözlərin tezlik göstəricilərindəki fərqlər aşkar olunmuşdur.

III fəslin **“M.Ə. Sabirin mətn korpusunun yaradılması”** adlı yarım-fəslində M.Ə. Sabirin mətnlər korpusunun yaradılması da təsvir edilir. Hazırda bu korpusda 274 satıra, 12 məktub, 4 hekayə və felyeton, 5 tərcümə, 22 uşaq şeir və hekayələri, 11 müxtəlif şeir, 11 qəzəl, 38 məqalə və müxbir qeydi, 37 təhmis vardır. M.Ə. Sabirin bədii mətnləri üzrə tezlik lüğətləri və konkordanslar tərtib edilmiş, həm ümumi korpus, həm də ayrı-ayrı əsərlər üzrə ən çox istifadə olunan vahidlər müəyyənləşdirilmişdir. M.Ə. Sabirin bütün əsərlərinin ümumi göstəriciləri aşağıdakı kimidir: word tokens – 67 138, word types – 21 361. Ən çox işlənən vahidlər isə bunlardır: “bir” – 1371 işlənmə; “bu” – 1037 işlənmə; “ki” – 861 işlənmə; “nə” – 632 işlənmə; “və” – 504 işlənmə; ilə – 476 işlənmə; “də” – 475 işlənmə; “hər” – 459 işlənmə; “da” – 304 işlənmə; “var” – 304 işlənmə; “mən” – 223 işlənmə. Bununla yanaşı, M.Ə. Sabirin dilinin bir sıra xüsusiyyətləri də üzə çıxarılmışdır. Məsələn, böyük satirikin dilində «məktəb» sözü “uşqol” formasında müxtəlif hallarda işlədilmişdir: “uşqol” – 1; “uşqola – 3; “uşqolada” – 1; “uşqoladan” – 1; “uşqolalar” – 1; “uşqolaların” – 1; “uşqolanın” – 2; “uşqolasına” – 1; “uşqolasına” – 2 və s.

**“M. Füzulinin (poetik) mətnlər korpusunun yaradılması”** adlı yarım-fəsilə isə M. Füzulinin mətnlər korpusunun yaradılmasından bəhs edilir. Hal-hazırda korpusa müəllifin 305 qəzəli,

“Leyli və Məcnun”, “Tiryək və şərab”, “Məyvələrin söhbəti”, “Şikayətnamə” poemaları daxildir. Tezlik lüğətləri və konkordanslar tərtib edilmiş, Füzulinin dil xüsusiyyətləri müəyyənləşdirilmiş, tokenlər və söz formalarının kəmiyyət nisbətləri həm ayrı-ayrı əsərlər, həm də ümumi mətn korpusu üzrə təhlil edilmişdir. Əsas nəticə isə “M. Füzulinin poetik korpusu”nun yaradılmasıdır. Burada 31 372 token və 9 289 söz forması qeydə alınmışdır. Həmçinin Füzuli poeziyasının dil xüsusiyyətləri üzə çıxarılmışdır: “etmək”, “eyləmək” və “olmaq” kimi köməkçi fellərin geniş işlənməsi, əvəzlilərin bağlayıcı funksiyasında istifadəsi, bəzi ərəb mənşəli sözlərin Azərbaycan dilinin felləri ilə birləşməsi yolu ilə mürəkkəb sözlərin yaradılması və s. Məsələn, yalnız “ol” sözünün işlənməsi 324 dəfə qeydə alınmış və tezliyinə görə dördüncü yeri tutmuşdur. «Olmaq» söz forması qəzəllərdə 16 dəfə, “olsam” – 12 dəfə, “olsun” – 13 dəfə, “olmuş” – 45 dəfə işlədilmişdir. “Et” söz forması “Divan”da 90 dəfə, “etdi” – 55 dəfə, “etdim” – 44 dəfə, “etdin” – 29 dəfə, “etmək” – 17 dəfə; “eylər” – 90 dəfə; “eylə” – 60 dəfə; “eyləmək” – 17 dəfə; “eylədi” – 15 dəfə; “eylədin” – 13 dəfə işlənməmişdir və s.

*Nəticədə* əsas çıxarışlar, eləcə də dissertasiyanın məzmunundan irəli gələn ümumiləşdirmələr təqdim edilir.

Tədqiqat Azərbaycan müəlliflərinin (şairlərin və yazıçıların) mətn korpuslarını, tədqiq olunan bədii mətnlərə daxil olan vahidlərin konkordanslarını, həmçinin bu vahidlərin tezlik siyahıları və lüğətlərini tərtib etmək məqsədilə həyata keçirilmişdir.

1. Mətn anlayışının çoxsahəli sistem kimi dəqiqləşdirilməsi; mətnin əsas kateqoriyaları (müttəfiqlik, bütövlük) müəyyən edilmiş və “mətn” və “diskurs” anlayışları ayrılmışdır;

2. Bədii mətnin xüsusi statusu müəyyən edilmiş, onun zəngin üslubi potensiala malik olması və geniş miqyaslı lingvostatistik və korpus tədqiqatlarının aparılmasına imkan verməsi təsdiqlənmişdir;

3. Korpus dilçiliyi və lingvostatikanın müstəqil elmi istiqamətlər kimi formalaşması izlənilmiş, mətn korpuslarının dilin tədqiqini asanlaşdırması, vahidlərin işlənmə qanunauyğunluqlarını və tendensiyalarını üzə çıxardığı nümayiş etdirilmişdir;

4. Dil tədqiqində lingvostatistika və korpus dilçiliyinin rolu əsaslandırılmış, Azərbaycan dilinin Milli Korpusunun yaradılmasının strateji əhəmiyyəti elmi cəhətdən əsaslandırılmışdır;

5. A.S. Puşkinin mətnlər korpusu materialı əsasında korpus metodlarının effektivliyi təsdiqlənmiş, onların Azərbaycan filologiyasında tətbiqi üçün geniş perspektivlər açır.

6. Korpus texnologiyalarının tətbiqi əhəmiyyəti müəyyən edilmiş, bunlara tezlik lüğətlərinin və konkordansların tərtibi, avtomatik mətn emalı, maşın tərcüməsi, müəlliflik atributikası və xarici dillərin tədrisində istifadəsi daxildir;

7. Xarici dillərin tədrisində korpus metodlarının tətbiqi çərçivəsində rus dilində ən çox işlədilən sözlərin Azərbaycan dilinə tərcüməsi üzrə lüğət hazırlanmışdır. Həmçinin isimlər, fellər, sifətlər və digər nitq hissələrinin müvafiq tərcümələri ilə birlikdə tezlik siyahıları tərtib edilmişdir ki, bu da onların tədris prosesində və metodik vəsaitlərin hazırlanmasında istifadəsini təmin edir;

8. N. Gəncəvi, M.Ə. Sabir və M. Füzulinin mətnlər korpusları yaradılmış; tezlik lüğətləri və konkordanslar tərtib edilmiş, onların tezlik və leksik-qrammatik xüsusiyyətləri, o cümlədən etnospesifik leksika, ərəb-fars mənşəli alınmalar və fərdi müəllif leksikası müəyyən edilmişdir;

9. Aşağı tezliyi olan leksika fərdi üslubun göstəricisi kimi çıxış edir; nadir rast gəlinən vahidlərin isə unikal sözləri, neologizmləri və idiolektizmləri əhatə etdiyi müəyyən edilmişdir;

10. A.S. Puşkinin “Kapitan qızı” povesti və onun Azərbaycan dilinə tərcüməsi, həmçinin N. Gəncəvinin “Leyli və Məcnun” poemasının rus və Azərbaycan dillərinə tərcümələrinin lingvostatistik təhlili aparılmış; situativ modallığın əsas predikatlarının tezlik göstəriciləri və onların orijinal və tərcümədəki fərqləri müəyyən edilmişdir;

11. M.Ə. Sabir və M. Füzulinin xarakterik dil xüsusiyyətləri müəyyən edilmiş və onların yaradıcılığının Azərbaycan ədəbi dilinin inkişafına təsiri təsdiqlənmişdir;

12. Bu tədqiqat Azərbaycanda korpus dilçiliyi sahəsində yenilikçi xarakter daşıyır: ilk dəfə sistemli şəkildə Azərbaycan müəlliflərinin

mətn korpusları və tezlik lüğətləri yaradılmış, bu isə Azərbaycan dilinin Milli Korpusunun formalaşması üçün əsas qoymuşdur.

***Dissertasiyanın əsas məzmunu müəllifin aşağıdakı elmi əsərlərində əksini tapmışdır:***

1. Ситуативная модальность и лексические средства ее выражения как объект корпусных исследований: [Электронный ресурс] / URL: <https://www.stephanos.ru/izd/2019/2019-36.pdf>

2. Компьютерная лингвистика и применение лингвостатистических методов при исследовании художественных текстов // – Ankara: Bilimsel araştırmalar, – 2021. – S. 33-35.

3. Особенности этноспецифической лексики на материале поэмы «Лейли и Меджнун» Низами Гянджеви: [Электронный ресурс] / – 2023. URL: <https://bsu-uni.edu.az/storage/konfranslar/azerbaycansunasligin-aktual-meseleleri-2023-2-cild-2.pdf>

4. Bədii mətn linqvostatistik tədqiqat obyektı kimi // – Bakı: II International congress of modern studies in social sciences and humanities «In Memory of Victory Day and Martyrs» konfrans materialları, – 2021. Vol. 1, – s. 248-255;

5. Azərbaycan yazıçılarının bədii mətn korpuslarının yaradılmasında süni intellektin metod və texnologiyaları: [Электронный ресурс] / – Bakı: “Azərbaycan və dünya ədəbiyyatşünaslığı: problemlər və perspektivlər” mövzusunda Respublika Konfransının materialları. URL: <https://baau.edu.az/az/announcement/azerbaycan-ve-dunya-edebiyatsunasligi-problemler-ve-perspektivler-movzusunda-respublika-konfrans-84>

6. Создание корпуса художественных текстов азербайджанского языка: [Электронный ресурс] / – М.: Материалы Международной научной конференции студентов, аспирантов и молодых учёных «Ломоносов-2025» – в Азербайджане. URL: <https://lomonosov->

msu.ru/archive/Lomonosov\_2025/data/9614/183797\_uid145271\_report.pdf

7. «Хамсе» Низами Гянджеви как объект корпусного и лингвостатистического исследования // – М.: Сборник статей LVII международной научно-практической конференции, – 2023. – с. 170-171.

8. Корпусное исследование на материале азербайджанского и русского языков. Создание корпуса текстов Низами Гянджеви // – М.: Текст. Исследования молодых учёных, – 2023. № 1–2 (8–9), с. 121-128.

9. Лингвостатистический анализ основных предикатов «ситуативной модальности на материале повести А.С. Пушкина «Капитанская дочка» и ее перевода на азербайджанский язык»: [Электронный ресурс] / – М.: Материалы Международного молодежного научного форума «Ломоносов-2022», – 2022. URL: [https://lomonosov-msu.ru/archive/Lomonosov\\_2022/data/25755/uid145271\\_5e4c9960aa04707225038b26eb9137a62141a505.doc](https://lomonosov-msu.ru/archive/Lomonosov_2022/data/25755/uid145271_5e4c9960aa04707225038b26eb9137a62141a505.doc)

10. Лингвостатистический анализ на материале перевода «Лейли и Меджнун» Низами Гянджеви»: [Электронный ресурс] / – М.: Материалы Международного молодежного научного форума «Ломоносов-2022». URL: [https://confhub.ru/archive/Lomonosov\\_2022/data/26116/146919\\_uid145271\\_report.pdf](https://confhub.ru/archive/Lomonosov_2022/data/26116/146919_uid145271_report.pdf)

11. Понятия «текст» и «художественный текст» в современной лингвистике // – Баки: Filologiya məsələləri, – 2023. № 12, – s. 196-203.

12. Проблема перевода. Трудности перевода теологической литературы на азербайджанский язык: [Электронный ресурс] / – Баку: “Müasir dünyada din təhsili” mövzusunda respublika elmi konfransının materialları. URL: <https://ait.edu.az/az/news/all/muasir-dunyada-din-t-hsili-movzusunda-respublika-elmi-konfransinin-materiallari-cap-olunub/>

13. Использование некоторых показателей возможности в корпусе текстов А.С. Пушкина: [Электронный ресурс] / – М.:

Материалы Международного молодежного научного форума «Ломоносов-2020». URL: [https://lomonosov-msu.ru/archive/Lomonosov\\_2020\\_2/data/section\\_36\\_19498.htm](https://lomonosov-msu.ru/archive/Lomonosov_2020_2/data/section_36_19498.htm)

14. Корпусное исследование сатир Мирза Алекпера Сабира. Создание корпуса текстов Сабира // *Filologiya məsələləri*, – Bakı: Elm və təhsil, – 2023. №14, – s. 425-432.

15. Лингвостатистика, или квантативная лингвистика: закон Ципфа, закон Мензерата – Альтманна преимущества лингвостатистических исследований // – *Gəncə: Elmi xəbərlər*, – 2023. № 4, s. 226-229.

16. Лингвостатистический анализ перевода «Лейли и Меджнун» Н. Гянджеви //– Bakı: II International congress of modern studies in social sciences and humanities «Year of Shusha-2022» konfrans materialları, – 2022. Vol. 3, – s. 207-211.

17. Создание поэтического корпуса газелей М. Физули // – М.: Балтийский гуманитарный журнал, – 2024. Т. 13. № 1(46), – с. 113-116

18. Развитие азербайджанской корпусной лингвистики. Создание корпуса текстов азербайджанских авторов (поэтов и писателей) и национального корпуса азербайджанского языка: [Электронный ресурс] / – Махачкала: Вестник Дагестанского государственного университета. Серия 2, Гуманитарные науки, – 2024. № 2. URL: <https://cyberleninka.ru/article/n/razvitie-azerbaydzhanskoy-korpusnoy-lingvistiki-sozdanie-korpusa-tekstov-azerbaydzhanskih-avtorov-poetov-i-pisateley-i>

19. Корпусная лингвистика: истоки зарождения, терминологический аппарат. Применение методов корпусной лингвистики при анализе текста // – Bakı: *Elmi əsərlər*, – 2023. Т. 17, № 4 (56), – с. 149-153.

20. Сопоставление морфологических категорий имени существительного в славянских и тюркских языках (на примере русского и азербайджанского языков): [Электронный ресурс] / URL:

[https://www.philol.msu.ru/~slavphil/books/shirokova\\_2024.pdf](https://www.philol.msu.ru/~slavphil/books/shirokova_2024.pdf)

Dissertasiyanın müdafiəsi 28 oktyabr 2025 tarixində saat 11:00 Azərbaycan Milli Elmlər Akademiyası Nəsimi adına Dilçilik İnstitutu nəzdində fəaliyyət göstərən ED 1.06 Dissertasiya şurasının iclasında keçiriləcək.

Ünvan: Bakı, AZ 1143, H.Cavid prospekti 115, V mərtəbə, AMEA Nəsimi adına Dilçilik İnstitutu.

Dissertasiya ilə Azərbaycan Milli Elmlər Akademiyası Nəsimi adına Dilçilik İnstitutunun kitabxanasında tanış olmaq mümkündür.

Avtoreferatın elektron versiyası Azərbaycan Milli Elmlər Akademiyası Nəsimi adına Dilçilik İnstitutunun rəsmi internet saytında yerləşdirilmişdir.

Avtoreferat 27 sentyabr 2025 tarixində zəruri ünvanlara göndərilmişdir.

Çapa imzalanıb: 25.09.2025

Kağız formatı: 60x84 1/16

Həcm: 43 103

Tiraj: 70