

<https://doi.org/10.62837/2025.8.132>

**QƏDİROVA GÜLÇÖHRƏ QƏDİR QIZI  
MİRƏZYEVA AYSU ATAMOĞLAN QIZI**

Azərbaycan Universiteti, Ceyhun Hacıbəyli, 71, Bakı, Azərbaycan  
e-mail: [gulchohra.gadirova@au.edu.az](mailto:gulchohra.gadirova@au.edu.az) [Aysu.Mirzayeva@au.edu.az](mailto:Aysu.Mirzayeva@au.edu.az)

**KORPUS DİLÇİLİYİNİN YARANMASI VƏ İNKİŞAFI: NƏZƏRİ VƏ  
PRAKTİKİ YANAŞMALAR**

**Xülasə**

Məqalədə korpus dilçiliyinin tarixi inkişafı, nəzəri əsasları və tətbiq sahələri araşdırılır. Korpuslar vasitəsilə dilin real istifadəsi empirik metodlarla təhlil olunur. Tədqiqatda texnoloji yeniliklərin, o cümlədən süni intellekt və multimodal korpusların rolu vurğulanır. Korpus dilçiliyinin metodoloji əhəmiyyəti və gələcək perspektivləri ümumiləşdirilir.

**Açar sözlər:** korpus dilçiliyi, empirik yanaşma, texnologiya, annotasiya, tətbiq sahələri

**1. GİRİŞ**

Korpus dilçiliyi dil tədqiqatında müasir dövrdə ən mühüm metodoloji nailiyyətlərdən biri hesab olunur. Onun yaranması və inkişafı dilin elmi təsvirinə, mətnlərin sistematik təhlilinə və dilçilikdə empirik yanaşmaların tətbiqinə yeni imkanlar açmışdır. Bu tədqiqatın məqsədi korpus dilçiliyinin yaranmasını, inkişaf mərhələlərini, nəzəri əsaslarını, alətlərini və texnoloji vasitələrini, tətbiq sahələrini, müasir tendensiyalarını və metodoloji təsirlərini əhatəli şəkildə araşdırmaqdır [1, səh. 15].

Korpus dilçiliyi dilin istifadəsinə əsaslanan empirik yanaşmanı təmsil edir. Onun əsasında geniş mətn kolleksiyalarının sistematik toplanması, təsnifatı və kompüter vasitəsilə təhlili dayanır. Bu yanaşma dilin strukturunu, funksiyalarını və dəyişməsinə real istifadə nümunələri əsasında öyrənməyə imkan verir [2, səh. 22].

Bu tədqiqat korpus dilçiliyinin həm nəzəri, həm də praktiki aspektlərini araşdıraraq dilçilik sahəsinə elmi töhfə verməyi hədəfləyir.

**2. TARİXİ KONTEKST**

Korpus dilçiliyinin tarixi XX əsrin ortalarına qədər uzanır. İlk korpus layihələrindən biri kompüter dəstəkli mətn təhlilinin əsasını qoyan Brown Korpusudur [3, səh. 7]. Sonrakı dövrdə Britaniya Milli Korpusu (BNC) və Müasir Amerika İngilis Dili Korpusu (COCA) kimi böyük korpuslar yaradılmışdır. Bu inkişaf korpus dilçiliyinin metodoloji imkanlarını genişləndirmiş və dil tədqiqatlarında empirik yanaşmanı daha da gücləndirmişdir.

Korpusların tarixi mərhələləri aşağıdakı kimi ümumiləşdirilə bilər:

- Əl ilə toplanmış kiçik korpuslar dövrü
- Kompüter dəstəkli korpusların inkişafı
- Müasir dövrdə veb-korpusların və multimodal korpusların yaranması

Bu mərhələlər korpus dilçiliyinin formalaşmasında texnologiyanın və metodologiyanın rolunu açıq şəkildə göstərir.

### **3. NƏZƏRİ ƏSASLAR**

Korpus dilçiliyi nəzəri olaraq dilin istifadəsinə əsaslanan empirik yanaşmaya söykənir. Bu yanaşmanın əsas prinsipləri dilin real istifadəsini öyrənmək, mətn nümunələrinə əsaslanmaq və statistik analiz metodlarını tətbiq etməkdir [1, səh. 30].

#### **1. Empirik Yanaşma**

Korpus dilçiliyi dilin təsviri üçün empirik məlumatların toplanmasını və analizini əsas götürür. Bu yanaşma nəzəri dilçiliyi praktiki təhlillə birləşdirərək dilin struktur və funksional xüsusiyyətlərini öyrənməyə imkan verir. Empirik yanaşma dilin statistik nümunələrinə əsaslanaraq dilin qaydalarını və istifadə meyllərini müəyyən etməyə çalışır [2, səh. 35].

#### **2. Teoriyalar və Modellər**

Korpus dilçiliyi müxtəlif nəzəri modellərlə inteqrasiya olunur. Bunlar arasında ən çox istifadə edilənlər:

- Frekans təhlili — dil vahidlərinin istifadəsinin tezliyini öyrənmək üçün.
- Kollokasiya təhlili — sözlərin birlikdə istifadəsinin modellərini müəyyən etmək üçün.
- Açar söz təhlili — korpusda xüsusi terminlərin və ifadələrin əhəmiyyətini analiz etmək üçün [4, səh. 45].

#### **3. Korpus Dilçiliyi və Dilçilik Nəzəriyyələri**

Korpus dilçiliyi dilçilik nəzəriyyələrinə yeni perspektivlər təqdim edir. Məsələn, funksional dilçilik korpus məlumatlarından istifadə edərək dilin kommunikativ funksiyalarını və kontekstual istifadəsini təhlil edir. Həmçinin, konstruksiya dilçilik (construction grammar) korpus dilçiliyinin empirik metodlarını nəzəri baza kimi istifadə edir [5, səh. 70].

### **4. KORPUS ALƏTLƏRİ VƏ TEXNOLOGİYASI**

Korpus dilçiliyi yalnız nəzəri baza ilə məhdudlaşmır, həm də texnoloji alətlərin inkişafı ilə sıx bağlıdır. Müasir korpus tədqiqatlarında istifadə olunan alətlər və texnologiyalar dilçilərə böyük mətn həcmində dəqiq və sistemli analiz aparmağa imkan verir.

#### **1. Korpus İdarəetmə Sistemləri**

Korpus idarəetmə sistemləri mətnlərin toplanması, saxlanması, təsnifatı və analizini asanlaşdırır. Bu sistemlər arasında ən geniş istifadə olunanlar:

- AntConc — frekans analizləri və kollokasiya təhlili üçün.
- WordSmith Tools — mətnlərin konkordans və tezlik analizləri üçün.
- Sketch Engine — korpusların geniş spektrdə təhlilinə imkan verən güclü platformadır [6, səh. 95].

#### **2. Annotasiya və Təsnifat**

Annotasiya (mətin işarələmə) prosesi korpus dilçiliyində vacib mərhələdir. Annotasiya mətnin morfoloji, sintaktik və semantik səviyyələrdə işarələnməsini təmin edir. Bu proses avtomatik və ya manual şəkildə həyata keçirilə bilər. Müasir annotasiya sistemləri NLP (Natural Language Processing) texnologiyaları ilə inteqrasiya olunur ki, bu da dilin daha dəqiq və sürətli təhlilini mümkün edir [7, səh. 125].

### **3. Multimodal Texnologiyalar**

Müasir korpus dilçiliyi yalnız mətnə əsaslanmır, həm də multimodal məlumatların təhlilini əhatə edir. Multimodal korpuslar audio, video, şəkil və jestləri də əhatə edir ki, bu da dilin kontekstual və funksional istifadəsini daha dərindən öyrənməyə imkan verir [8, səh. 160].

### **4. Bulud Texnologiyaları və Böyük Məlumatlar**

Bulud texnologiyaları və "big data" yanaşmaları korpus dilçiliyinə yeni imkanlar gətirmişdir. Böyük korpuslar (məsələn, COCA və Google Books Ngram Corpus) milyardlarla söz məlumatını ehtiva edir ki, bu da dilin diaxronik dəyişməsini və qlobal dil nümunələrini araşdırmağa imkan verir [9, səh. 200].

### **5. Gələcək Texnoloji Perspektivlər**

Korpus dilçiliyinin gələcək inkişafı süni intellekt (AI) və maşın öyrənməsi texnologiyalarının korpus tədqiqatlarına inteqrasiyasına bağlıdır. Bu istiqamətdə inkişaflar annotasiya keyfiyyətini artıracaq, multimodal korpusların təhlilini genişləndirəcək və real vaxtda korpus yaradılmasına imkan verəcəkdir [7, səh. 305].

## **5. TƏTBİQLƏR**

Korpus dilçiliyi yalnız nəzəri və metodoloji sahədə deyil, həm də praktiki tətbiqlərdə geniş istifadə olunur. Onun imkanları dil tədrisi, lüğətşünaslıq, tərcümə, diskurs təhlili, sosio-dilçilik və forensik dilçilik kimi sahələrdə əhəmiyyətli rol oynayır.

### **1. Lüğətşünaslıq və Lüğət Hazırlığı**

Korpus dilçiliyi lüğətşünaslıqda sözlərin istifadəsini, tezliklərini və kollokasiya nümunələrini təhlil etmək üçün geniş istifadə olunur. Bu məlumatlar lüğətlərin hazırlanmasında əsas mənbə kimi xidmət edir [10, səh. 88]. Məsələn, korpuslardan əldə olunan data əsasında sözlərin mənalari, istifadə kontekstləri və nümunə cümlələri daha dəqiq müəyyən edilir.

### **2. Dil Tədrisi**

Korpus dilçiliyi dil tədrisində müəllimlərə və tələbələrə real mətn nümunələri təqdim edərək dilin təbii istifadəsini öyrətmək imkanı verir. Learner corpora (tələbə korpusları) vasitəsilə dil öyrənənlərin səhvləri və istifadə nümunələri analiz edilərək tədris materialları hazırlanır [11, səh. 54]. Bu yanaşma dil tədrisində empirik əsaslı metodologiyanın tətbiqinə imkan yaradır.

### **3. Tərcümə Araşdırmaları**

Korpus metodları tərcümə prosesində dilin struktur və üslub xüsusiyyətlərini analiz etmək üçün geniş istifadə olunur. Tərcümə korpusları müxtəlif dillərdə paralel mətnləri əhatə edərək tərcümə strategiyalarının və nümunələrinin araşdırılmasına şərait yaradır [4, səh. 79].

#### **4. Diskurs Təhlili**

Diskurs təhlilində korpus dilçiliyi geniş miqyasda mətnlərin strukturunu və dil modellərini təhlil etməyə imkan verir. Media diskursu, siyasi nitq, elmi mətnlər kimi sahələrdə korpus analizləri diskursun struktur və funksional xüsusiyyətlərini açır [12, səh. 110].

#### **5. Sosio-Dilçilik və Forensik Dilçilik**

Sosio-dilçilikdə korpuslar dilin sosial kontekstlərdə istifadəsini araşdırmaq üçün istifadə olunur. Forensik dilçilikdə isə hüquqi mətnlər və danışmaq materiallarının korpus təhlili hüquqi məsələlərdə sübut kimi istifadə olunur [13, səh. 142].

#### **NƏTİCƏ**

Bu tədqiqat korpus dilçiliyinin yaranması, inkişaf mərhələləri, nəzəri və metodoloji əsasları, texnoloji imkanları və tətbiq sahələrini əhatəli şəkildə araşdırmışdır. Korpus dilçiliyi müasir dilçilikdə empirik yanaşmanın əsas dayaqlarından biri kimi çıxış edir və dilin real istifadəsinə əsaslanan təhlillərlə elmi araşdırmaların obyektivliyini və etibarlılığını artırır. Tədqiqatın nəticələri göstərir ki, korpusların dizaynı, annotasiya prosesləri və statistik analiz metodları dilin struktur və funksional xüsusiyyətlərinin daha dərinə öyrənilməsinə imkan yaradır. Nəzəri baxımdan bu yanaşma dil modellərinin və konseptual çərçivələrin formalaşmasına töhfə verir. Praktiki sahələrdə isə korpus metodologiyası lüğətşünaslıq, tərcümə, dil tədrisi, diskurs və forensik təhlil kimi müxtəlif istiqamətlərdə effektiv şəkildə tətbiq olunur. Müasir tendensiyalar, o cümlədən süni intellekt, maşın öyrənməsi, multimodal və çoxdilli korpuslar sahənin inkişafını sürətləndirməklə, dilçilikdə yeni araşdırma imkanları yaradır. Bu texnoloji və metodoloji yeniliklər korpus dilçiliyinin yalnız indiki deyil, həm də gələcək dil tədqiqatlarında əsas rol oynayacağını göstərir.

#### **ƏDƏBİYYAT**

1. Atkins, B. T. S., & Rundell, M. (2008). *The Oxford Guide to Practical Lexicography*. Oxford University Press.
2. Baker, M. (1993). *Corpus Linguistics and Translation Studies: Implications and Applications*. In M. Baker, G. Francis, & E. Tognini-Bonelli (Eds.), *Text and Technology* (pp. 233-258). John Benjamins.
3. Baker, P., Gabrielatos, C., & McEnery, T. (2013). *Discourse Analysis and Media Attitudes: The Representation of Islam in the British Press*. Cambridge University Press.
4. Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge University Press.

5. Coulthard, M., & Johnson, A. (2007). *An Introduction to Forensic Linguistics: Language in Evidence*. Routledge.
6. Francis, W. N., & Kučera, H. (1964). *Brown Corpus Manual*. Brown University.
7. Goldberg, A. E. (2006). *Constructions at Work: The Nature of Generalization in Language*. Oxford University Press.
8. Granger, S. (2009). *Learner Corpora*. In A. Lüdeling & M. Kytö (Eds.), *Corpus Linguistics: An International Handbook* (pp. 259-275). Mouton de Gruyter.
9. Hunston, S. (2002). *Corpus in Applied Linguistics*. Cambridge University Press.
10. Jurafsky, D., & Martin, J. H. (2021). *Speech and Language Processing* (3rd ed. draft). Available at <https://web.stanford.edu/~jurafsky/slp3/>
11. Kilgarriff, A., et al. (2014). *The Sketch Engine: Ten Years On*. *Lexicography*, 1(1), 7-36.
12. Michel, J.-B., et al. (2011). *Quantitative Analysis of Culture Using Millions of Digitized Books*. *Science*, 331(6014), 176-182.
13. O'Halloran, K. L., Tan, S., & Smith, B. A. (2016). *Multimodal Corpus Linguistics: A New Frontier*. Routledge.

**Gadirova G.**

**Mirazyeva A.**

## **THE EMERGENCE AND DEVELOPMENT OF CORPUS LINGUISTICS: THEORETICAL AND PRACTICAL APPROACHES**

### **Abstract**

The paper examines the development, theory, and applications of corpus linguistics. It highlights how real language use is analyzed through empirical methods. The role of new technologies, including AI and multimodal corpora, is discussed. The methodological importance and future outlook of corpus linguistics are summarized.

**Keywords:** corpus linguistics, empirical approach, technology, annotation, applications

**Гадирова Г.**

**Миразиева А.**

## **ВОЗНИКНОВЕНИЕ И РАЗВИТИЕ КОРПУСНОЙ ЛИНГВИСТИКИ: ТЕОРЕТИЧЕСКИЕ И ПРАКТИЧЕСКИЕ ПОДХОДЫ**

### **Резюме**

В статье рассматривается история, теоретические основы и области применения корпусной лингвистики. Анализируется использование эмпирических методов для изучения языка на основе реальных текстов. Отдельное внимание уделяется роли новых технологий, включая

искусственный интеллект и мультимодальные корпуса. Обобщается методологическая значимость и перспективы развития корпусной лингвистики.

**Ключевые слова:** корпусная лингвистика, эмпирический подход, технологии, аннотирование, применение

**Rəyçi: f.e.d., professor, Azad Yəhya oğlu Məmmədov**