

Linqvostatistik metodlarla lüğətlərin tərtibi sahəsi

Nigar Səxavət qızı Vəlizadə
AMEA-nın Nəsimi adına Dilçilik
İnstitutunun kiçik elmi işçisi
E-mail: n.valizade@mail.ru

Rəyçilər: filol.ü.e.d., prof. M.Ə. Mahmudov,
filol.ü.e.d., prof. K.A. Vəliyeva

Açar sözlər: lüğət, bədii mətn, linqvostatistik təhlil, təhlil metodu

Key words: dictionary, literary text, linguostatistical analysis, method of analysis

Ключевые слова: словарь, художественный текст, лингвостатистический анализ, метод анализа

Azərbaycan dilinin materialları əsasında linqvostatistik metodlarla lüğətlərin tərtibi sahəsində artıq müəyyən işlər görülmüşdür. Qeyd etmək lazımdır ki, Azərbaycan dilinin tezlik, əks lüğətləri ilə yanaşı abidələrin (Kitabi-Dədə Qorqud) və klassiklərin (Məhəmməd Füzuli) dilinin linqvostatistik aspektdə tədqiqi ilə bağlı da araşdırmalar aparılmışdır. Həmin lüğətlər daha çox dilin leksik qatını əhatə etsə də, mətnlərin öyrənilməsi ilə bağlı da istifadə oluna bilərlər.

İlkin təcrübə kimi həmin lüğətlər və onların bədii mətnin leksik baxımdan tədqiqində istifadə imkanları ilə tanış olaq.

1. “Azərbaycan qəzet dilinin tezlik lüğəti” (tərtibçilər – K. Vəliyeva, M. Mahmudov, V. Pines) Azərbaycan qəzet dili materialları əsasında tezlik lüğətinin tərtibi sahəsində ilk təcrübədir. Lüğətdə leksik vahidlər işlənmə tezlikləri ilə birlidə qeyd olunmuşdur. Bu tipli lüğətlər leksikanın statistik modeli kimi leksikologiyanın və leksikoqrafiyanın bir sıra nəzəri məsələlərini təhlil etmək imkanı yaradır. Qeyd etmək lazımdır ki, dilin leksik tərkibinin statistik tədqiqi nəticəsində əldə olunmuş məlumatlardan tipoloji təsnifatlarda, funksional və müəllif üslublarının müəyyənləşdirilməsində, stenoqrafiyada, tədrisdə, məşin tərcüməsində, elektron lüğətlərin tərtibində, eləcə də nitq mədəniyyətinə aid araşdırmalarda istifadə oluna bilər. Haqqında danışılan lüğət ötən əsrin 70-ci illərində “BESM-2” elektron hesablama məşinlərinin köməyi ilə tərtib olunmuş, əlifba-tezlik və tezlik siyahılarından ibarətdir (2).

2. “Azərbaycan dilinin tezlik lüğəti (söz kökləri) I cild (tərtibçilər – M. Mahmudov, Ə. Fətullayev, S. Məmmədova və b.). Lüğətin annotasiyasında göstəriləyi kimi təqdim olunan tezlik lüğətinin tərtibi üçün Azərbaycan dilinin bütün üslubları təmsil olunmaqla təxminən 50 milyona yaxın söz-formadan ibarət mətnlər linqvostatistik təhlilə cəlb olunmuşdur. Elmi, bədii, publisist, rəsmi, məişət və s. üslublara dair materialların elektron variantlarının mövcudluğu tədqiqatın hərtərəfli olmasına və statistik seçmə baxımından etibarlılığına imkan yaratmışdır. Lüğətdə təxminən 12 minə yaxın söz kökü əhatə olunmuşdur. Lüğət–minimumların, danışq kitabçalarının hazırlanması, tədris materialının seçimi prosesində haqqında söhbət gedən tezlik lüğətindən faydalanmaq olar. Dilin sözmələgətirmə tiplərinin, sözdəyişdirmə imkanlarının araşdırılması. orfoqrafiya və qrafikanın təkmilləşdirilməsi baxımından da bu lüğətin böyük səmərəsi ola bilər (3).

3. “Məhəmməd Füzulinin nəzm əsərlərinin əlifba-tezlik sözlüyü” (tərtibçilər – K. Vəliyeva, M. Mahmudov, C. Rəhmanov, V. Sultanov). Kitabda görkəmli Azərbaycan şairi Məhəmməd

məd Füzulinin anadilli nəzm əsərlərinin əlifba – tezlik sözlüyü verilmişdir. Azərbaycan dilinin tarixinin öyrənilməsi baxımından araşdırma olduqca əhəmiyyətli mənbə kimi istifadə olunabilir (7).

4. “Kitabi Dədə Qoqqud”un statistik təhlili (tərtibçilər K. Vəliyeva, M. Mahmudov, V. Pines, C. Rəhmanov, V. Sultanov). Kitabda dünya türkologiya elmində ilk dəfə olaraq Azərbaycan xalqının qədim yazılı abidəsi “Kitabi-Dədə Qorqud” dastanlarının statistik tədqiqinin nəticələri əksini tapmışdır. Kompüterlərin köməyi ilə alınmış siyahılar abidənin dili haqqında bir çox qiymətli nəticələr əldə etməyə imkan verir. Azərbaycan dilinin tarixini öyrənmək baxımından araşdırma olduqca əhəmiyyətlidir (4).

6. “Azərbaycan dilinin əks əlifba lüğəti” (tərtibçilər M. Mahmudov, Ə. Fətullayev). Azərbaycan dilinin əks əlifba lüğəti orfoqrafiya lüğətinin sözlüyü əsasında tərtib olunmuşdur və təxminən 80 minə yaxın sözü əhatə edir. Adi lüğətlərdən fərqli olaraq əks lüğətlərdə sözlər başlanğıcdan sona doğru deyil, əksinə sondan başlanğıca doğru əlifba sırasına düzülür. Adi lüğətlərdə eyni başlanğıca malik sözlər, əks lüğətlərdə isə eyni sonluğa malik sözlər eyni qrupda yerləşir. Annotasiyada qeyd olunduğu kimi Azərbaycan dilçiliyində ilk təcrübə olan bu lüğətdən nəzəri dilçiliyin bir sıra məsələləri ilə əlaqədar, maşın tərcüməsi və mətnlərin formal təhlili sistemlərində, tekstologiya və paleoqrafiyada, əlyazmaların oxunuşu və bərpası işlərində, naməlum yazı sistemlərinin deşifrə olunması prosesində və qafiyə lüğəti kimi istifadə olunabilir (1).

Yuxarıda qeyd etdiyimiz lüğətlərdə ən çox işlənən leksik vahidlərin bəzi ümumi xüsusiyyətlərini nəzərdən keçirək.

Azərbaycan dilinin tezlik lüğətində yüksək tezlikli 50 söz:

Və, ol(maq), bu, ed(et(mək)), bir, ki, gör(mək), ilə, il, ver(mək), o,keç(mək), baş, de(mək), iş, üçün, bil(mək), öz, çox, dövlət, biz, ölkə, mən, kim, gəl(mək), gün, isə, yer, onlar, məlumat, sonra, çıx(maq), on, nazir, ara, prezident, göstər(mək), əsas, tərəf, art(maq), al(maq), yeni, böyü(mək), respublika, belə, bildir(mək), hər, qarşı, təşkil, qeyd.

Azərbaycan qəzet dilinin tezlik lüğətində yüksək tezlikli 50 söz:

Və, et(mək), bu, bir, yüz, o, ol(maq), il, partiya, ilə, sovet, min, ki, Azərbaycan, iş, üçün, doqquz, respublika, iyirmi, xalq, təsərrüfat, yer, çox, on, də, böyük, iki, da, sov.ikp, rayon, daha, məhsul, qurultay, bütün, beşillik, sahə, ver(mək), əmək, kənd, hər, ölkə, plan, kommunist, beş, mk, baş, dövlət, gün, öz, yeni.

“Kitabi-Dədə Qorqud”un statistik lüğətində yüksək tezlikli 50 söz:

De(mək), ay(dır), bir, qara, qazan, nə, gəl(mək), ağ, xanım, bu, ol, oğul, mərə, dəxi, oğuz, yigit, kim, ol(maq), soyla(maq), mən, mana, gör(mək), at, ala, al(maq), oğlan, sən, kafər, bəg, üzəri(nə), var(maq), var, qız, san(a), qırq, ver(mək), gün, qoca, üç, bura, dəli, ər, söylə(mək), xəbər, görk, yer, qalın, yoq, adam, ilə.

Məhəmməd Füzulinin azərbaycandilli nəzm əsərlərində yüksək tezlikli 50 söz:

Kim, ki, bir, ol, bu, ilə, ey, hər, nə, ol(maq), eylə(mək), mən, var, həm, gər, yenə, et(mək), o, yox, tək, gül, qıl(maq), idi, can., çün(üçün), çox, eşq, gün, ta, min, içrə, q(ğ)əm, dəm, yar, və, gör(mək), göz, üzrə, könül, degil, hiç, xoş, necə, özgə, yüz, qan, nə, ah, ver(mək), ikən (6).

Bu tezlik göstəriciləri mətnin leksik quruluşunun tədqiqi baxımından əhəmiyyətlidir. Qeyd etmək lazımdır ki, bəzən ayrı-ayrı leksik vahidlər də bütövlükdə müəllif üslubunun müəyyənləşdirilməsində mühüm rol oynaya bilər. Yazıçı üslubunda bəzən müəyyən qəlib ifadələr,

məhəlli sözlər, dialekt və şivə ünsürləri az və ya çox dərəcədə öz əksini tapa bilər ki, bu da müəllif üslubunun müəyyənləşdirilməsi işində nəzərə alınmalıdır. Əlbəttə, onu da xüsusi vurğulamaq lazımdır ki, müəllif üslubunun müəyyənləşdirilməsi sintaktik vahidlər səviyyəsində daha qabarıq müşahidə olunur. Söz birləşmələri, sadə cümlə, yarımçıq və natamam cümlə tipləri, mürəkkəb cümlə, inversiya və digər yardımçı vasitələr bu prosesdə daha önəmli hesab olunur.

Lüğətlərin bir növü olan tezlik lüğətləri mətn daxilindəki sözlərin (söz formaları, söz birləşmələri) nömrəli şəkildə tətbiqi əsasında qurulur. Bu lüğətləri istənilən dildə sözlərin istifadəsinin ədədi xüsusiyyətlərini verir. Lüğətləri, lüğətin və mətnin strukturundakı ədədi nümunələri müqayisə etməyə imkan verir.

Azərbaycan dilinin tezlik lüğətlərinin tərtibində elmi, bədii, publisist, rəsmi, məişət və s. üslublar nəzərə alınaraq tərtib olunmalı və bütün söz–forma linqvostatistik baxımından kompüterdə işlənməli, rəqəmlər, simvollar, təsadüfi sözlər və ya xüsusi adlar və s. burada öz əksini tapmalıdır.

Azərbaycan dilinin təxminən 90 illik lüğətcilik tarixi var. Azərbaycan dilçiliyinin bir çox sahələrinə aid qiymətli tədqiqatlar aparmış A.M. Qurbanov lüğətlərin əhəmiyyətindən danışaraq yazır: “Lüğətlər xalqın mədəni inkişafı və tələbatı əsasında meydana çıxmışdır. Belə ki, xalqın mədəniyyəti, dünyagörüşü artdıqca onun tələblərini ödəmək üçün müəyyən vasitəyə ehtiyac duyulur” (5).

Məqalənin aktuallığı. Azərbaycan dilinin tezlik, əks lüğətləri ilə yanaşı abidələrin və klassiklərin dilinin linqvostatistik aspektdə tədqiqi ilə bağlı da araşdırmalar aparılmışdır. Həmin lüğətlər daha çox dilin leksik qatını əhatə etsə də, mətnlərin öyrənilməsi ilə bağlı da istifadə oluna bilər... Mətnin bədii təhlili ənənəvi lüğətlərlə yanaşı, statistik lüğətlərdən də müəyyən ideya və fikirləri əxz etməyə istiqamətlənməlidir. Belə olduqda mətnin təhlili daha dolğun, elmi, hərtərəfli və məntiqi olacaqdır.

Məqalənin elmi yeniliyi. Yeni təlim metodlarından istifadə etməklə daha maraqlı və səmərəli iş hazırlanmışdır. Nəzəri biliyin daha çox praktik əhəmiyyətindən danışılmış, strukturu və xüsusi metodları barədə məlumat verilmişdir.

Məqalənin praktik əhəmiyyəti və tətbiqi. Dilin potensial imkanlarından istifadə üsulları, formalarının öyrənilməsi baxımından böyük əhəmiyyət kəsb edir. Ali təhsil müəssisələrində, magistratura və ya doktorantura səviyyəsində olan tələbələr üçün əlavə vəsait kimi istifadə oluna bilər.

Ədəbiyyat

1. Azərbaycan dilinin əks əlifba lüğəti (tərtibçilər: M. Mahmudov, Ə. Fətullayev). Bakı: Nurlan, 2004, 524 s.

2. Azərbaycan qəzet dilinin tezlik lüğəti (tərtibçilər: K. Vəliyeva, M. Mahmudov, V. Pines). Bakı: Nurlan, 2004, 264 s.

3. Azərbaycan dilinin tezlik lüğəti (söz kökləri). I cild (tərtibçilər: M.Ə. Mahmudov, Ə.B. Fətullayev və b.). Bak: Elm, 2010, 464 s.

4. “Kitabi-Dədə Qorqud”un statistik təhlili (tərtibçilər K. Vəliyeva, M. Mahmudov, V. Pines, C. Rəhmanov, V. Sultanov). Bakı: Elm, 1999, 248 s.

5. Qurbanov A. Azərbaycan dilçiliyi problemləri. I cild. Bakı: Nurlan, 2004, 444 s.

6. Mahmudov M. Ə. Kompüter dilçiliyi. Bakı: Elm və təhsil, 2013, 356 s.

7. Məhəmməd Füzulinin nəzm əsərlərinin əlifba-tezlik sözlüyü (tərtibçilər: K. Vəliyeva, M. Mahmudov, C. Rəhmanov, V. Sultanov). Bakı: Elm, 2004, 548 s.

Н.С. Велизаде

**Сфера составления словарей
лингвостатистическими методами**

Резюме

Вышеуказанное акцентирует важность научного значения изучения языка писателя математико-стилистическими методами. Очень полезно применение лингвостатистических методов при всестороннем изучении языка писателя и письменных памятников (источников) в азербайджанской литературе и культуре. Используя опыт и теоретические знания полученные в этой области в дальнейшем было бы полезно усилить исследования в области лингвостатистического изучения языка и стиля писателя, а так же памятников письменности.

N.S. Valizadeh

**The field of compiling dictionaries
by linguostatistical methods**

Summary

The above is important in terms of once again highlighting the scientific significance of the study of the writer's language by mathematical and statistical methods. The application of linguostatistical methods in the comprehensive study of the writer's language and written monuments in Azerbaijani literature and culture is very useful. It would be useful to intensify research in the field of linguostatistical study of the writer's language and style, as well as written language monuments, using the experience and theoretical knowledge gained in this field.

Redaksiyaya daxil olub: 24.08.2021