

Azərbaycan dilində mətnlərin oxunabilirlik səviyyəsinin qiymətləndirilməsi metodları

İsmayıl Sadıqov

İnformasiya Texnologiyaları İnstitutu, Bakı, Azərbaycan
ismayil.sadigov@gmail.com

Xülasə — Məqalədə ingilis dilindəki mətnlər üçün oxunabilirlik düsturlarının Azərbaycan dilindəki mətnlərə tətbiqi probleminə toxunulur. Bu problemin həlli üçün düsturlardakı orta cümlə uzunluğu, orta söz uzunluğu kimi parametrlərə uyğun əmsalların korrektə edilməsindən, eləcə də “tanış sözlər” siyahısının hazırlanmasından danışılır. İngilis dilindəki mətnlər üçün dörd oxunabilirlik düsturunun Azərbaycan dilindəki mətnlər üçün uyğunlaşdırılmış variantı təqdim olunur.

Açar sözlər — *mətnin mürəkkəbliyi; oxunabilirlik düsturları; Fleşin oxuma asanlıığı düsturu; Fleş-Kinkeydin təhsil səviyyəsi düsturu; Deyl-Çellin yeni oxunabilirlik düsturu; Speşin oxunabilirlik düsturu.*

I. GİRİŞ

İnsanlar xüsusən məktəb illərində qazandıqları biliklərin böyük bir hissəsini dərslərdən alırlar. Tədqiqatlara görə şagirdlərdə oxu vərdişlərinin inkişafını ləngidən əsas amillərdən biri dərslərin mətnlərinin mürəkkəbliyi, daha dəqiqi, mətnlərin mürəkkəblilik dərəcəsi ilə şagirdlərin yaş səviyyəsi arasında uyğunsuzluqdur. Həddən artıq mürəkkəb

mətnləri şagirdlər başa düşmədən əzbərləməyə məcbur olurlar.

Mətnin mürəkkəbliyi mətnin obyektiv xassəsi olub, onu oxuyan insandan asılı deyil. Mətndəki komponentlərin qiymətini, məsələn, cümlələrin uzunluğunu, mücərrəd sözlərin nisbətini və başqa parametrləri dəyişməklə onun mürəkkəbliyini də dəyişmək olar.

Mətnin mürəkkəbliyi onun başa düşülməsini çətinləşdirir. *Mətnin çətinliyi* təkcə onun mürəkkəbliyindən deyil, həm də oxucunun hazırlıq səviyyəsindən asılıdır. Bir (hazırlıqlı) oxucu üçün asan olan mətn, başqa bir (hazırlıqsız) oxucu üçün çətin ola bilər.

II. MƏTNİN MÜRƏKKƏBLİK FAKTORLARININ AŞKARLANMASI

Mətnin mürəkkəbliyinə təsir edən faktorları aşkarlamaq üçün belə bir sadə üsuldan istifadə olunur: Ekspert qiymətləndirməsi əsasında sadə və ya mürəkkəb hesab edilən müəyyən sayda mətn götürülür. Bu mətnlər ayrı-ayrı göstəricilər – məntiqi quruluş, mövzu, cümlələrin uzunluğu və başqa parametrlər üzrə müqayisə edilir. Əgər göstəricinin qiyməti asan mətnlərdən mürəkkəb mətnlərə keçdikcə dəyişilirsə, demək, bu

göstərici mətnin mürəkkəbliyinə faktorlarından biridir. Məsələn, əgər sadə və mürəkkəb mətnlərin müqayisəsi zamanı mürəkkəb mətnlərdə daha çox naməlum sözlərin və uzun cümlələrin olması müəyyən olunursa, deməli, mətnin mürəkkəbliyi sözlərin tanışı olmasından və cümlələrin uzunluğundan asılıdır.

Tədqiqatlar nəticəsində mətnin mürəkkəbliyinə ən çox təsir edən faktorlar müəyyən edilmişdir. Onlardan bəziləri aşağıdakı cədvəldə göstərilmişdir (Cədvəl 1).

Cədvəl 1. Mətnlərin mürəkkəbliyinə təsir edən faktorlar

№	Parametrin (faktorun) adı
1	Orta cümlə uzunluğu (sözlərlə) – Average sentence length in words
2	Orta cümlə uzunluğu (hecalarla) – Average sentence length in syllables
3	Sadə cümlələrin sayı – Number of simple sentences
4	Bir abzasdakı cümlələrin sayı – Number of sentences per paragraph
5	Birhecalı sözlərin faizi – Percentage of monosyllables
6	Çoxhecalı sözlərin faizi – Percentage of polysyllables
7	Fərqli sözlərin faizi – Percentage of different words
8	Birinci, ikinci və üçüncü şəxs əvəzlərinin sayı – Number of first, second, and third-person pronouns

III. “OXUNABİLİRLİK” ANLAYIŞI VƏ OXUNABİLİRLİK DÜSTURLARI

Mətnin mürəkkəbliyinə səviyyəsini göstərmək üçün “*oxunabilirlik*” (ingilis

dilində “*readability*”) terminindən istifadə olunur.

Hazırda hər hansı mətnin, xüsusən də şagirdlərin yaş xüsusiyyətlərini nəzərə almaqla tədris mətnlərinin oxunabilirliyini qiymətləndirmək üçün bir sıra riyazi modellər – *oxunabilirlik düsturları* mövcuddur. Elmi və populyar ədəbiyyatda *oxunabilirlik düsturları (readability formulas)* ifadəsi ilə yanaşı, *oxunabilirlik testləri (readability tests)*, *oxunabilirlik metrikası (readability metrics)* ifadələrindən də istifadə olunur.

Oxunabilirlik sahəsində ilk elmi tədqiqatlara 1920-ci illərdə ABŞ-da başlanılmış və ötən yüz ildə 100-lərlə oxunabilirlik düsturu işlənib hazırlanmışdır. Təbii ki, bu düsturların hamısı populyarlıq qazana bilməmişdir. Ancaq onlardan bəziləri bugün də geniş istifadə olunur:

- Fleşin oxuma asanlıığı düsturu (Flesch Reading Ease Formula);
- Fleş-Kinkeydin təhsil səviyyəsi düsturu (Flesch-Kincaid Grade Level Formula);
- Deyl-Çellin oxunabilirlik düsturu (Dale-Chall Readability Formula);
- Speşin oxunabilirlik düsturu (Spache Readability Formula);
- SMOG oxunabilirlik düsturu (SMOG Readability Formula);
- Frayn oxunabilirlik düsturu (Fry Readability Formula);
- Qanningin duman indeksi (Gunning Fog Index);
- FORCAST oxunabilirlik düsturu (FORCAST Readability Formula);

IV. ORTA CÜMLƏ UZUNLUĞU VƏ ORTA SÖZ UZUNLUĞUNA ƏSASLANAN OXUNABİLİRLİK DÜSTURLARI

Oxunabilirlik düsturları içərisində ən geniş istifadə olunanı *Fleşin oxuma asanlığı düsturudur*. Bu düstur iki dəyişənə (faktora) – mətndəki cümlələrin orta uzunluğuna və sözlərdəki hecaların orta sayına əsaslanır:

$$K_{en} = 206.835 - (1.015 \times S) - (84.6 \times W), \quad (1)$$

burada K_{en} – mətnin mürəkkəbliyinin (oxuma asanlığının) göstəricisi, S – orta cümlə uzunluğu (sözlərin sayının cümlələrin sayına nisbəti), W isə sözdəki hecaların orta sayıdır (hecaların sayının sözlərin sayına nisbəti). [1]

Düsturun asan tətbiqi üçün belə bir metodika təklif olunur: hər hansı mətndən 100 söz götürülür, cümlələrin sözlərlə orta uzunluğu (orta cümlə uzunluğu) və sözlərin hecalarla orta uzunluğu (orta heca sayı) hesablanır. Oxuma asanlığının qiyməti 100-dən 0-dək dəyişə bilər (Cədvəl 2). Ancaq hazırda çoxlu sayda proqram təminatı olduğundan mətnin ölçüsünün böyük götürülməsi hesablama işlərini çətinləşdirmir.

Cədvəl 2. Fleşin oxuma asanlığı qiymətləri

Oxuma asanlığının qiyməti (K_{en})	Təsviri
0 – 29	çox çətin mətn
30 – 49	çətin mətn
50 – 59	nisbətən çətin mətn
60 – 69	standart mətn
70 – 79	nisbətən asan mətn
80 – 89	asan mətn
90 – 100	çox asan mətn

Geniş istifadə olunan başqa bir oxunabilirlik düsturu – *Fleş-Kinkeydin təhsil səviyyəsi düsturudur*. Əslində, bu düstur Fleşin oxuma asanlığı düsturunun təkmilləşdirilmiş variantıdır. Bu düsturun tətbiqindən alınan qiymət mətnin hansı təhsil səviyyəsi üçün nəzərdə tutulduğunu göstərir. Düstur aşağıdakı kimidir:

$$GL_{en} = (0.39 \times S) + (11.8 \times W) - 15.59, \quad (2)$$

burada GL_{en} uyğun təhsil səviyyəsini göstərir. Məsələn, 8.3 qiyməti mətnin orta səviyyəli 8-ci sinif şagirdi (adətən, 12–14 yaş) üçün anlaşılacaq olacağını bildirir. Düsturdan göründüyü kimi, cümlələrdəki sözlərin sayı və sözlərdəki hecaların sayı nə qədər az olarsa, alınan qiymət də bir o qədər kiçik olar, yəni mətni daha aşağı yaşlı oxucular başa düşə bilər.

V. MƏTNDƏKİ TANIŞ OLMAYAN SÖZLƏRİN SAYINA ƏSASLANAN OXUNABİLİRLİK DÜSTURLARI

1948-ci ildə ABŞ maarifçisi Edqar Deyl və psixoloq, yazıçı Janna Çell hazırda çox populyar və etibarlı hesab edilən oxunabilirlik düsturunu təklif etmişlər. [2] *Deyl-Çellin oxunabilirlik düsturunda* orta cümlə uzunluğu və mürəkkəb sözlərin faizindən istifadə olunmuşdur. Burada “mürəkkəb sözlər” dedikdə müəlliflər tərəfindən tərtib olunmuş və 80%-i 4-cü sinif səviyyəsində oxucuya məlum olan 3000 sadə sözdən ibarət siyahıya daxil olmayan sözlər nəzərdə tutulur. Əsas ideya ondan ibarətdir ki, oxucu tanış sözləri daha asan qəbul edir və başa düşür. Deyl-Çellin oxunabilirlik düsturu Fleşin oxuma asanlığı

düsturunun bəzi çatışmazlıqlarını aradan qaldırmaq məqsədilə işlənib hazırlanmışdır.

1995-ci ildə E.Deyl və J.Çell 3000 sözlük siyahını yenidən işləmiş və öz düsturlarının yeni versiyasını dərc etdirmişlər. [3] *Deyl-Çellin yeni oxunabilirlik düsturu* özündən əvvəlki düsturların üstün cəhətlərini nəzərə almış və mətnin oxunabilirliyini daha dəqiqliklə qiymətləndirir. Düstur aşağıdakı kimidir:

$$Raw\ Score = 0.1579 \times PDW + 0.496 \times ASL, \quad (3)$$

burada *Raw Score* – ilkin (xam) bal; *PDW* – mürəkkəb sözlərin (Deyl-Çellin siyahısında olmayan sözlərin) faizi; *ASL* – orta cümlə uzunluğudur.

Əgər *PDW* 5%-dən yüksək olarsa, yekun balı almaq üçün ilkin balın üzərinə 3.6365 əlavə olunur, yəni *Adjusted Score = Raw Score + 3.6365*, əks halda, *Adjusted Score = Raw Score*. Düsturun tətbiqi nəticəsində alınan qiymətin təhsil səviyyəsinə uyğunluğu aşağıdakı cədvəldə göstərilmişdir (Cədvəl 3).

Cədvəl 3. Deyl-Çell düsturundan alınan qiymətin təhsil səviyyəsinə uyğunluq cədvəli

Yekun bal	Təhsil səviyyəsi
4.9 və aşağı	orta səviyyəli 4-cü və ondan aşağı sinif şagirdi asanca başa düşür
5.0–5.9	orta səviyyəli 5-ci və ya 6-cı sinif şagirdi asanca başa düşür
6.0–6.9	orta səviyyəli 7-ci və ya 8-ci sinif şagirdi asanca başa düşür
7.0–7.9	orta səviyyəli 9-cu və ya 10-cu sinif şagirdi asanca başa düşür
8.0–8.9	orta səviyyəli 11-ci və ya 12-ci sinif şagirdi asanca başa düşür
9.0–9.9	orta səviyyəli kollec tələbəsi asanca başa düşür
10.0 və yuxarı	orta səviyyəli kollec məzunu asanca başa düşür

Deyl-Çellin yeni oxunabilirlik düsturuna bənzəyən başqa bir düstur 1953-cü ildə Corc Speş təklif etmişdir. Bu düsturda da məntdəki orta cümlə uzunluğu və “mürəkkəb” sözlərin faizi kimi parametrlərdən istifadə olunmuşdur. *Speşin oxunabilirlik düsturu* adlandırılan bu düstur ibtidai təhsil səviyyəsi (1–4-cü siniflər) üçün mətnlərin oxunabilirliyini qiymətləndirmək üçün nəzərdə tutulmuşdur. [4]

$$Index\ Spache = 0.141 \times ASL + 0.086 \times PDW + 0.839, \quad (4)$$

burada *PDW* – mürəkkəb (tanış olmayan) sözlərin faizi; *ASL* – orta cümlə uzunluğudur.

C.Speş 3000 sözlük siyahı əvəzinə, öz siyahısını təklif etmişdir: Speşin siyahısı ibtidai sinif şagirdlərinin bildiyi təxminən 1000 sözdən ibarətdir və bu siyahı ABŞ məktəblərinin 3-cü sinif şagirdləri ilə aparılan eksperiment nəticəsində müəyyən olunmuşdur: şagirdlərin 80%-nin qeyd etdiyi söz *tanış sözlər siyahısına* daxil edilmişdir.

VI. OXUNABİLİRLİK DÜSTURLARIN MODİFİKASIYA EDİLMƏSİ

Nəzərə almaq lazımdır ki, bu düsturların hamısı təcrübi yolla ingilis dilində olan mətnlər üçün alınmışdır. Hər bir dilin özünəməxsus sintaksisi, yəni söz və cümlə quruluşu vardır. Məsələn, ingilis dilində sözlərin orta uzunluğu Azərbaycan, yaxud rus dili ilə müqayisədə qısadır. Bundan başqa, müxtəlif dillərdə sözlərdə olan hecaların miqdarı fərqli qaydada hesablanır: dillərin əksəriyyətində sözdəki hecaların

sayı həmin sözdəki saıtların miqdarı ilə müəyyən olunsada, bəzi dillərdə bu say tələffüzdən asılı olur. Məsələn, ingilis dilində “analyse” sözü üç hecadan ibarətdir, ancaq “analyses” sözü dörd hecalıdır. Ona görə də, bir dildəki mətnlər üçün hazırlanmış oxunabilirlik düsturunu başqa dildəki mətnlərə tətbiq etdikdə düzgün olmayan nəticələr alınabilir. Daha dəqiq desək, ingilis dilinin sintaksisinə uyğun hazırlanmış oxunabilirlik düsturlarını Azərbaycan dilində olan mətnlərə tətbiq etmək üçün həmin düsturlardakı əmsallar korrektə edilməlidir.

Tədqiqat göstərmişdir ki, ingilis dilindəki mətnlər Azərbaycan dili ilə müqayisədə orta hesabla 0.77 cümlə uzundur, ingilis dilindəki sözlər isə Azərbaycan dilindəki sözlərdən 1.91 heca qıtsadır. [5] Yəni Fleşin oxuma asanlığı düsturunu Azərbaycan dilindəki mətnlərə tətbiq etmək üçün orta cümlə uzunluğunun əmsalı (S) 0.77 dəfə, hecalarla orta söz uzunluğunun (W) əmsalı isə 1.91 dəfə korrektə edilməlidir. Nəticədə *Azərbaycan dilindəki mətnlər üçün Fleşin oxuma asanlığı düsturu* aşağıdakı şəkildə düşür:

$$K_{az} = 206.835 - (1.318 \times S) - (44.3 \times W). \quad (5)$$

Eyni qayda ilə Fleş-Kinkeyd düsturu da Azərbaycan dili mətnləri üçün uyğunlaşdırılmışdır. Beləliklə, *Azərbaycan dilindəki mətnlər üçün Fleş-Kinkeydin təhsil səviyyəsi düsturu* aşağıdakı şəkildə olmuşdur: [6]

$$GL_{az} = (0.51 \times S) + (6.18 \times W) - 15.59 \quad (6)$$

Azərbaycan dilindəki mətnlər üçün modifikasiya olunmuş növbəti *eyl-Çellin*

yeni oxunabilirlik düsturu (3) və *Speşin oxunabilirlik düsturudur* (4). Bu düsturlardakı orta cümlə uzunluğu (ASL) parametrinin əmsalının korrektə edilməsi ilə bağlı problem yoxdur. Bu korrektəni etdikdən sonra *Azərbaycan dilindəki mətnlər üçün Deyl-Çellin yeni oxunabilirlik düsturu* aşağıdakı şəkildə düşəcəkdir:

Əgər $PDW > 5\%$ isə, onda:
 $Adjusted\ Score = Raw\ Score + 3.6365,$ (7)
əks halda:
 $Adjusted\ Score = Raw\ Score.$

Eyni qayda ilə Speşin oxunabilirlik düsturu uyğunlaşdırıldıqdan sonra *Azərbaycan dilindəki mətnlər üçün Speşin oxunabilirlik düsturu* aşağıdakı şəkildə olacaqdır: [7]

$$GL = 0.183 \times ASL + 0.086 \times PDW + 0.839 \quad (8)$$

Əsas problem (7) və (8) düsturlarını Azərbaycan dilindəki mətnlərə tətbiq etmək üçün PDW kəmiyyətinin qiymətini, yəni mürəkkəb sözlərin faizini hesablamaqdadır. Çünki bunun üçün E.Deyl və J.Çell tərəfindən ingilis dili üçün hazırlanmış 3000 sözlük siyahının, eləcə də C.Speş tərəfindən təklif olunmuş ibtidai sinif şagirdlərinin bildiyi təxminən 1000 sözdən ibarət siyahının Azərbaycan dili üçün bənzərlərinin olması vacibdir. Ancaq Azərbaycan dili üçün indiyədək bənzər siyahılar olmadığından məqalə müəllifi tərəfindən belə siyahılar da hazırlanmışdır.

Azərbaycan dilindəki mətnlərin mürəkkəblik səviyyəsini qiymətləndirmək üçün məqalə müəllifi tərəfindən iki proqram təminatı da işlənib hazırlanmışdır.

Bunlardan biri “*Mətn analizi*” tətbiqi proqramı, digəri isə *www.oxunabilir.az* veb-proqramıdır. [8]

NƏTİCƏ

- Digər oxunabilirlik düsturlarının Azərbaycan dili üçün modifikasiya edilməsindən ötrü müvafiq proqram təminatının olması vacibdir.
- Belə proqram təminatının hazırlanmasına maddi dəstəyin olması lazımdır.
- Kompüterlərdə hazırlanmış və çap olunmuş bütün kitablardakı mətnlərin nəşriyyatlar tərəfindən “xam” şəkildə vahid mərkəzə təhvil verilməsi və tədqiqatçılar üçün həmin resursların əlçatanlığının təmin edilməsi vacibdir.

İSTİNADLAR

[1] Flesch R. Marks of a readable style. Columbia University contributions to education, 1943, no. 897. New York: Bureau of Publications, Teachers College, Columbia University.

- [2] Dale E., Chall J.S. A formula for predicting readability. Educational research bulletin. 1948, Jan. 21 and Feb. 17, 27: 1-20, 37-54.
- [3] Chall J.S., Dale E. Readability revisited, the new Dale-Chall readability formula. Cambridge, MA: Brookline Books, 1995.
- [4] Spache G. A new readability formula for primary-grade reading materials // Elementary school journal, 1953, 53: pp.410-413.
- [5] Sadıqov İ.C. Azərbaycan dili mətnlərinin mürəkkəbliyinin qiymətləndirilməsi üçün modifikasiya olunmuş Fleş düsturu // "İnformasiya texnologiyaları problemləri" jurnalı, 2018, №1, s.46-58.
- [6] Sadıqov İ.J. Mathematical and information models for evaluating readability of texts in Azerbaijani language // El-Cezeri Journal of Science and Engineering, V.5 – N.3, 2018, pp. 888-903.
- [7] Sadıqov İ.C. Azərbaycan türkçesindeki metinlər için kelime sıklıklarına dayalı okunabilirlik formülleri // “Uluslararası Mühendislik ve Doğa Bilimleri Çalışmaları Kongresi”, Ankara, 07-09 may 2021.
- [8] Sadıqov İ.C. Azərbaycan dili mətnlərinin mürəkkəbliyinin qiymətləndirilməsi üçün modifikasiya olunmuş Fleş düsturu // "İnformasiya texnologiyaları problemləri" jurnalı, 2018, №1, s.46-58.