

ÜÇ VƏ DAHA ÇOX FAKTORLU OXUNABİLİRLİK DÜSTURLARININ QURULMASI

İsmayıl SADIQOV *

Annotasiya: Məqalədə mətnlərin mürəkkəblik dərəcəsini qiymətləndirmək üçün mövcud oxunabilirlik düsturları, onlarda istifadə olunan parametrlər (faktorlar) əsasında çoxparametrlilik oxunabilirlik düsturlarının yaradılması metodikası verilir. Bu məqsədlə çoxdəyişənli reqressiya modelinin qurulması, reqressorların və reqressiya əmsallarının müəyyənəndirilməsi yolu təsvir olunur. Mövcud oxunabilirlik düsturlarının təklif olunan çoxparametrlilik düsturunun xüsusi halları olduğu nümunələrlə göstərilir.

Açar sözlər: *Mətnin Mürəkkəbliyi, Oxunabilirlik Düsturları, Çoxdəyişənli Xətti Reqressiya Modeli, Qiymətləndirmə Şkalası, Çoxparametrlilik Oxunabilirlik Düsturları.*

Construction of Readability Formulas With Three and More Factors

Abstract: The article provides a methodology for creating multiparameter readability formulas based on existing readability formulas, parameters (factors) used in them to assess the degree of complexity of texts. For this purpose, the construction of a multivariable regression model, the way of determining regressors and regression coefficients is described. It is shown by examples that existing readability formulas are special cases of the proposed multiparameter readability formula.

Key Words: *Text Complexity, Readability Formulas, Multiple Linear Regression Model, Rating Scale, Multiparameter Readability Formulas.*

Giriş

Mətnin mürəkkəblik səviyyəsini, onun qavranılma dərəcəsini göstərmək üçün "oxunabilirlik" (ingiliscə: "readability") terminindən istifadə olunur. Oxunabilirliyin qiymətləndirilməsi üçün bir çox riyazi modellər qurulmuşdur ki, onların da əksəriyyəti riyazi düsturlar şəklindədir. İngilis dilindəki mətnlər üçün yüzlərlə *oxunabilirlik düsturu* işlənilib hazırlansa da, onların yalnız çox az qismi populyarlıq qazana bilmişdir [1].

Azərbaycan dilindəki mətnlər üçün ilk oxunabilirlik düsturları ingilis dilindəki mətnlər üçün hazırlanmış iki məşhur düsturun modifikasiya edilməsi nəticəsində alınmışdır. Bunlar Fleşin oxuma asanlıığı və Fleş-Kinkeydin təhsil səviyyəsi düsturlarıdır [2].

Problemin Qoyuluşu

Məlumdur ki, oxunabilirlik düsturlarının mütləq əksəriyyətində 1-2 parametrdən istifadə olunur. Hər parametr mətnin oxunabilirliyini (mürəkkəbliyini) bir yöndən qiymətləndirir. Buna görə də eyni bir mətn üçün hətta etibarlı hesab edilən müxtəlif oxunabilirlik düsturları fərqli nəticələr verə bilər. Təbii sual yaranır: düsturlardakı parametrlərin sayını artırmaqla daha hərtərəfli, daha "dəqiq" qiymətləndirmələrə nail olmaq mümkün deyilmi? Əlbəttə, mümkündür. Onda ikinci sual yaranır: bəs nə üçün belə çoxparametrlilik oxunabilirlik düsturları yaranmır?

Məsələ ondadır ki, "oxunabilirlik düsturları bumu" ötən əsrin ortalarında başlamış və 1980-ci illərdə davam etmişdir. [1] Bu düsturların tətbiqi alqoritmlərindən də aydın olur ki, onlar "əllə" hesablama üçün nəzərdə tutulub. Belə ki, hesablamaların asanlaşdırılması və hesablama müddətinin çox olmaması üçün bir çox düsturların alqoritmlərində qiymətləndirilən mətndən 100 sözdən ibarət fraqmentin götürülməsi qeyd olunub. Ancaq günümüzə fərdi kompüterlərin geniş istifadə olunduğu və mətnlərin analizi üçün zəruri proqram təminatının yaradılması problem olmadığından çoxparametrlilik oxunabilirlik düsturlarının yaradılmasına üçün ciddi maneə olmadığını düşünürük. Bunun üçün, sadəcə, iki məsələni həll etmək lazımdır: 1) yaradılacaq çoxparametrlilik oxunabilirlik düsturlarına mətnlərin mürəkkəbliyinə (oxunabilirliyinə) təsir edən hansı faktorlar daxil edilməlidir, 2) düsturlara daxil edilən parametrlərin əmsalları necə müəyyən edilməlidir?

Problemin Həlli Metodikası

Qoyulan problemin həll üçün belə bir metodika təklif edilir:

- 1) Ən geniş istifadə olunan oxunabilirlik düsturlarını götürmək;
- 2) Həmin düsturların hər birində istifadə olunan parametrləri (faktorları) qeyd etmək;
- 3) Hər bir parametrin istifadə tezliyini, yəni neçə oxunabilirlik düsturunda istifadə edildiyini müəyyənəndirmək və istifadə tezliyi daha yüksək olan bir neçə (məsələn, 4-5) parametri seçmək və bu parametrlər əsasında reqressiya tənliyi qurmaq;

* Ögr. Gör. Dr., Necmettin Erbakan Üniversitesi, AMYO, Adalet Bölümü, ayumutlu@erbakan.edu.tr, ORCID: 0000-0001-9500-5338

- 4) Qiymətləndirmə şkalasını müəyyənləşdirmək;
- 5) Reqrəssiya tənliyindən alınan qiymətlərin şkalaya düşməsi üçün parametrlərin əmsallarını müəyyənləşdirmək.

1) 1920-ci ildən 1980-ci ilədək ingilis dilindəki mətnlər üçün 200-dən artıq oxunabilirlik düsturu işlənib hazırlanmışdır. Təbii ki, bu düsturların hamısı populyarlıq qazana bilməmişdir, ancaq bəzi düsturlar bu gün də geniş istifadə olunur: [3]

- Fleş'in oxuma asanlıığı düsturu (Flesch Reading Ease Formula);
- Fleş-Kinkeyd'in təhsil səviyyəsi düsturu (Flesch-Kincaid Grade Level Formula);
- Fray'in oxunabilirlik düsturu (Fry Readability Formula);
- Qanning'in dumanlıq indeksi (Gunning Fog Index);
- Deyl-Çell'in oxunabilirlik düsturu (Dale-Chall Readability Formula);
- SMOG oxunabilirlik düsturu (SMOG Readability Formula);
- Speş'in oxunabilirlik düsturu (Spache Readability Formula);
- Pauers-Samner-Körl oxunabilirlik düsturu (Powers-Sumner-Kearl Readability Formula);
- FORCAST oxunabilirlik düsturu (FORCAST Readability Formula);
- Avtomatlaşdırılmış oxunabilirlik indeksi (Automated Readability Index).

Mövcud oxunabilirlik düsturlarının əksəriyyətinin əsasında xətti reqrəssiya modeli dayanır və bu modelin dəyişənləri kimi mətnin statistik parametrləri çıxış edir:

$$f(x, b) = b_0 + \sum_{i=1}^k b_i x_i$$

burada b_i – reqrəssiya əmsalları; x_i – reqrəssorlar (modelin mürəkkəblik faktorları); k – modelin faktorlarının sayıdır.

Oxunabilirlik düsturlarında, adətən, leksika və sintaksis ilə əlaqəli 1–2 parametrdən istifadə edilmişdir. Əsas tətbiq sahələri tədris mətnlərin qiymətləndirilməsi olan bu düsturlarda reqrəssiya əmsalları elə seçilmişdir ki, alınan nəticələr təklif olunan mətni başa düşməsi üçün oxucunun hansı təhsil səviyyəsinə malik olmasını və ya hansı yaşda olmasını göstərsin.

2) Göstərilən 10 oxunabilirlik düsturunun hər birində hansı faktorlardan istifadə olduğunu cədvəl şəklində göstərək (cədvəl 1).

Cədvəl 1. Oxunabilirlik düsturlarında istifadə olunan parametrlər

№	Oxunabilirlik düsturu	İstifadə olunan parametrlər
1	Fleş'in oxuma asanlıığı düsturu	ASL, ASW
2	Fleş-Kinkeyd'in təhsil səviyyəsi düsturu	ASL, ASW
3	Fray'in oxunabilirlik düsturu	ASL, ASW
4	Qanning'in dumanlıq indeksi	ASL, NHW
5	Deyl-Çell'in oxunabilirlik düsturu	ASL, PDW
6	SMOG oxunabilirlik düsturu	NST, NHW
7	Speş'in oxunabilirlik düsturu	ASL, PDW
8	Pauers-Samner-Körl oxunabilirlik düsturu	ASL, NSL
9	FORCAST oxunabilirlik düsturu	NSW
10	Avtomatlaşdırılmış oxunabilirlik indeksi	ASL, AWL

Burada:

ASL – average sentence length – orta cümlə uzunluğu (sözlərlə),

ASW – average number of syllables per word – orta söz uzunluğu (hecalarla),

AWL – average word length in symbols – orta söz uzunluğu (simvollarla),

NHW – hard words (number of words of more than two syllables) – çətin sözlərin sayı

(3 və daha çox hecalı sözlərin sayı),

NST – number of sentences – cümlələrin sayı,

NSL – number of syllables – hecaların sayı,

NSW – number of single-syllable words in a 150-word sample – 150 sözdən ibarət mətndə birhecalı sözlərin sayı,

PDW – percentage of difficult words (words not on the Dale-Chall word list) – çətin sözlərin faizi (Deyl-Çell siyahısında olmayan sözlər).

3) Bu parametrlər arasından ən çox istifadə olunan 4 parametri seçdirək. Cədvəldən görüldüyü kimi, adıçəkilən 10 ən populyar düsturun demək olar ki, hamısında orta cümlə uzunluğu (*ASL*) parametri iştirak edir. Qalan parametrlər isə bunlardır: orta söz uzunluğu (*ASW*), çətin sözlərin faizi (*PDW*), çoxhecalı sözlərin sayı (*NHW*). Bu parametrləri əsas götürsək, yuxarıdakı reqressiya modeli belə olacaq:

$$f(x, b) = b_0 + b_1 \times ASL + b_2 \times ASW + b_3 \times PDW + b_4 \times NHW$$

Yuxarıda adları çəkilən oxunabilirlik düsturlarının, demək olar ki, hamısına bu düsturun xüsusi halları kimi baxmaq olar. Məsələn, $b_0 = 206.835$, $b_1 = -1.015$, $b_2 = -84.6$, $b_3 = b_4 = b_5 = 0$ olduqda ingilis dilindəki mətnlər üçün Fleş'in oxuma asanlıığı düsturu alınır:

$$RE = 206.835 - 1.015 \times ASL - 84.6 \times ASW.$$

Yeri gəlmişkən, Azərbaycan dilindəki mətnlər üçün modifikasiya olunmuş bu düsturda müvafiq əmsallar $b_0 = 206.835$, $b_1 = -1.318$, $b_2 = -44.3$ kimi [4], rus dilindəki mətnlər üçün isə $b_0 = 206.836$, $b_1 = -1.3$, $b_2 = -60.1$ kimi [5] olacaqdır.

Reqressiya əmsallarının $b_0 = -15.59$, $b_1 = 0.39$, $b_2 = 11.8$, $b_3 = b_4 = b_5 = 0$ qiymətlərində Fleş-Kinkeyd'in təhsil səviyyəsi düsturu (ingilis dilindəki mətnlər üçün)

$$GL = 0.39 \times ASL + 11.8 \times ASW - 15.59;$$

$b_0 = 3.6365$ (əgər x_3 qiyməti 5%-dən yüksək olarsa, əks halda $b_0 = 0$), $b_1 = 0.496$, $b_3 = 0.1579$, $b_2 = b_4 = b_5 = 0$ qiymətlərində isə Deyl-Çell'in oxunabilirlik düsturu:

$$Adjusted\ Score = \begin{cases} 0.1579 \times PDW + 0.496 \times ASL + 3.6365, & PDW > 5 \\ 0.1579 \times PDW + 0.496 \times ASL, & PDW \leq 5 \end{cases}$$

alınır.

4) Bundan sonra qiymətləndirmə şkalasını müəyyənləşdirmək və reqressiya modeindən alınan qiymətlərin şkalaya düşməsi üçün parametrlərin əmsallarını müəyyənləşdirmək qalır. Qiymətləndirmə şkalası olaraq R.Fleşin 1949-cu ildə təklif etdiyi şkalanı götürmək olar (cədvəl 2). Həmin bölgüyə görə, 0–30 aralığında qiymət alan mətnlər “çox çətin”, 30–50 aralığında – “çətin”, 50–60 aralığında – “nisbətən çətin”, 60–70 aralığında – “standart”, 70–80 aralığında – “nisbətən asan”, 80–90 aralığında – “asan”, 90–100 aralığında qiymət alan mətnlər isə “nisbətən asan” hesab olunur. [6]

Cədvəl 2. Fleşin oxuma asanlıığı qiymətləri

Oxuma asanlıığının qiyməti	Təsviri	Uyğun olduğu təhsil səviyyəsi
0–30	Çox çətin	Kollec məzunu
30–50	Çətin	13–16-cı siniflər
50–60	Nisbətən çətin	10–12-ci siniflər

60–70	Standart	8 və 9-cu sinif
70–80	Nisbətən asan	7-ci sinif
80–90	Asan	6-cı sinif
90–100	Çox asan	5-ci sinif

5) Parametrlərin əmsallarını müəyyənləşdirmək üçün isə belə bir metodikadan istifadə olunmasını məqsəduyğun hesab edirik: seçilmiş oxunabilirlik düsturları üçün orada istifadə olunan parametrlərin mümkün olan müxtəlif qiymətlərində mətnin mürəkkəblik dərəcəsinin qiymətlərini hesablamaq lazımdır. Məsələn, Fleşin oxuma asanlıığı düsturunda orta cümlə uzunluğu və orta söz uzunluğu parametrlərinin mümkün olan müxtəlif qiymətləri üçün uyğun oxuma asanlıığının qiymətləri aşağıdakı cədvəldə verilmişdir (cədvəl 3).

Cədvəl 3. Fleş düsturunda oxuma asanlıığı əmsalının orta cümlə uzunluğu və orta söz uzunluğundan asılılığı

Orta söz uzunluğu (ASW)	Orta cümlə uzunluğu (ASL)								
	5	6	7	8	9	10	11	12	
1.2	100	99	98	97	96	95	94	93	
1.3	92	91	90	89	88	87	86	85	
1.4	83	82	81	80	79	78	77	76	
1.5	75	74	73	72	71	70	69	68	
1.6	66	65	64	63	62	61	60	59	
1.7	58	57	56	55	54	53	52	51	
1.8	49	48	47	46	45	44	43	42	
1.9	41	40	39	38	37	36	35	34	
2.0	33	32	31	30	29	27	26	25	
2.1	24	23	22	21	20	19	18	17	

Analoji cədvəlləri başqa oxunabilirlik düsturları üçün də hazırlayıb, sonra ASL, ASW, PDW, NHW parametrlərinin hamısının iştirak etdiyi müxtəlif hallarda oxuma asanlıığının ekspert qiymətləndirilməsini aparmaq lazımdır. Bu qiymətləndirmənin nəticəsi olaraq oxuma asanlıığının göstərilən 4 parametrdən asılılıq funksiyasını (trend cizgisini) müəyyənləşdirmək olar.

Nəticə

Azərbaycan və Türkiyə, ortak tarixi və kültürel mirasa sahip bölgelerin korunması ve tanıtımı için işbirliğinde de soydaşlarının desteklerini alabilir. Bu sayede soydaşların ortak geçmişlerine daha fazla vurgu yapılabilir. Azərbaycan və Türkiyə arasındaki sivil toplum kuruluşları, ortak projeler yürüterek sosyal, kültürel ve ekonomik alanda işbirliği yapabilirler. Bu projeler, halklar arasında doğrudan etkileşimi artırabilir ve soydaşlık bağlarını güçlendirebilir.

KAYNAKÇA

1. DuBay W.H. The Classic Readability Studies // Costa Mesa, California: Impact Information, 2006.
 2. Sadıqov İ.C. Azərbaycan dili mətnlərinin mürəkkəbliyinin qiymətləndirilməsi üçün modifikasiya olunmuş Fleş düsturu // "İnformasiya texnologiyaları problemləri" jurnalı, 2018, №1, s.46–58.
 3. DuBay W.H. The Principles of Readability. Costa Mesa, California: Impact Information, 2004.
 4. Sadıgov I.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.
 5. Оборнева И.В. Автоматизированная оценка сложности учебных текстов на основе статистических параметров. Дис. канд. пед. наук, Москва: 2006.
- Алгулиев Р.М., Садыгов И.Дж. Оценивание удобочитаемости учебников на азербайджанском языке // "ScienceRise", 2018, №11(52), s.50–57