

YAPAY ZEKÂYA DAYALI OKUNABİLİRLİK FORMÜLÜNE DOĞRU BİR ADIM

A STEP TOWARD THE FORMULA OF READABILITY
BASED ON ARTIFICIAL INTELLIGENCE

Araştırma Makalesi

Ahmet BENZER¹

Makale gönderim tarihi 23 Mart 2020

Makale kabul tarihi 30 Nisan 2020

Özet

Yapay Zekâya dayalı bir okunabilirlik formülü ile yalnızca bir metnin okunabilirliğini değil aynı zamanda anlaşılabilirliğini de aynı anda tespit etmek mümkündür. Geleneksel okunabilirlik formüllerinde yalnızca metne dayalı özelliklerden hareketle düzey tespiti yapılırken yapay zekâya dayalı formüllerde metne dayalı daha fazla değişkenin yanı sıra okuyucu özelliklerinin de dâhil olduğu bir düzey ataması yapılmaktadır. Türkçe için en son geliştirilen formül 2010 yılında yapılmıştır. Aradan geçen 10 boyunca yeni bir formül geliştirilmemiştir. Bununla birlikte yapay zekâya dayalı formül çalışmaları dünyada 2009 yılı itibarıyla geliştirilmeye başlamıştır. Bu araştırmada parmakla, elle ya da gözle sayım yapılan okunabilirlik formüllerinden, Türkçe yapay zekâya dayalı okunabilirlik formülünün oluşmasında ilk basamağı oluşturmaya dönük Türkçenin yapısal özelliklerini ortaya koymak amaçlanmıştır. Araştırma, tarama modelinde hazırlanmıştır. Mevcut okunabilirlik çalışmalarının tamamı geleneksel yaklaşımla oluşturulmuş olup formüllerin neredeyse tamamında nicel değişkenler (kelime, cümle uzunluğu gibi) üzerine kurulduğu nitel değişkenlerin yani anlaşılabilirlik kısmının ise göz ardı edildiği tespit edilmiştir. Araştırma kapsamında belirlenen okunabilirlik formüllerinin tamamı 5, 6, 7 ve 8. sınıf Türkçe ders kitaplarındaki 12 metin ve 6 çocuk kitabı üzerinde uygulanmıştır. Araştırma sonucunda formüllerin düzey sonuçları arasında uyumsuzluk olduğu, Ateşman'ın (1997) formülüne göre kolay olarak belirlenen bir metin, Fleisch'te (1948) çok zor; Dale ve Chall'de (1948) 10-12 yaş olarak belirlenen bir metin ARI'da (1967) 16-17 yaş olarak belirlenmiştir. Ayrıca bu formüllerdeki yönerge eksikliği ise var olan değişkenlerin de doğru olarak tespit edilmesinde belirsizlikleri arttırmıştır. Okuyucuya dair değişkenlerin ise hiçbir formüllerde işe koşulmadığı tespit edilmiştir. Türkçenin yapısal özellikleri bağlamında en çok "cümle" başlığında farklı kullanımlar olduğu ortaya konmuştur. Araştırmada daha fazla eser üzerinden Türkçenin yapısal özelliklerinin tespit edilmesi, okunabilirliğin nicel değişkenlerle birlikte nitel değişkenlerin işe koşulduğu ve sürekli kendini güncelleyebilen yapay zekâya dayalı bir formülün hazırlanması önerilmiştir.

Anahtar Kelimeler: Okunabilirlik, yapay zekâ, anlaşılabilirlik.

Abstract

With a readability formula based on AI, it is possible to detect not only the readability of a text, but also its comprehensibility at the same time. In traditional readability formulas, only level determination is made based on text-based features, while artificial intelligence-based formulas are assigned a level that includes more text-based variables as well as reader features. The most recently developed formula for Turkish was made in 2010. No new formulas have been developed during the intervening 10. In addition, formula studies based on artificial intelligence started to be developed around the world, as of 2009. In this study, it was aimed to reveal the structural features of Turkish to form the first step in the formation of the readability formula based on Turkish artificial intelligence, one of the readability formulas counted by finger, hand or eye. The research was prepared in the screening model. All the existing readability studies have been created with a traditional approach, and it has been determined that almost all the formulas are based on quantitative variables (such as word, sentence length), and

¹ Doç. Dr., Marmara Üniversitesi Atatürk Eğitim Fakültesi Türkçe Eğitimi Bölümü Göztepe Kadıköy/İstanbul 34722. ahmetbenzer@gmail.com. ORCID: <https://orcid.org/0000-0003-3579-3699>

qualitative variables, namely comprehensibility, are ignored. All readability formulas determined within the scope of the research were applied on 12 texts and 6 children's books in 5, 6, 7 and 8th-grade Turkish textbooks. A text that is incompatible between the results of the formulas and easily determined according to the formula of Ateşman (1997) is very difficult in Flesch (1948); A text set in Dale and Chall (1948) as 10-12 years old were identified as 16-17 years in ARI (1967). Besides, the lack of instructions in these formulas increased the uncertainties in the correct determination of the existing variables. It was determined that variables related to the reader were not used in any formulas. In the context of the structural features of Turkish, it was revealed that there are different uses in the title "sentence". As a result of the research, it was proposed to determine the structural features of Turkish over more works, and to prepare a formula based on artificial intelligence where the readability and qualitative variables together with the quantitative variables were run, and that could update itself continuously.

Key Words: Readability, artificial intelligence, comprehensibility.

GİRİŞ

Yapay zekâ, 21. yüzyılın en önemli çalışmalarından biridir. İnsan gibi düşünebilen makineler, fikriyle ortaya çıkan bu çalışmalar zamanla hayatın her alanında kullanılmaya başlanmıştır. Sistemin en önemli özelliği, problemi tespit eden, değişimlere göre kendini güncelleyebilen, sonuçları kişiye veya zamana göre değişmeyen bir mantığına dayanmasıdır. "Bu sistemi Türkçe alanında okunabilirlik bağlamında nasıl kullanabiliriz?" düşüncesi bu çalışmanın var olma sebebidir. Bir sistem kurmak söz konusu ise ilk aşamada her bir değişkenin bütün versiyonlarının tespit edilmesi önem taşımaktadır. Huawei mühendisi Qun (2019), konuşma özelinde dile dair beklentilerden dolayı bilgisayar sistemlerinde eşi benzeri görülmemiş bir büyüme olduğunu belirtir ancak doğal dil işleme çalışmalarının henüz daha başlangıç seviyesinde olduğunu da ekler. Qun, sorunun anlamsal olduğunu yapay zekânın insanları anlamadığını belirtir ve akıllı cihazlara "restaurant bul" dediğimizde ciddi bir veri sağladığını ancak "açım" dediğimizde cihazın sonuç vermediği belirtir. Bu durumda gösteriyor ki bilgisayar sisteminde değişken olarak ne tanımlanırsa bilgisayar da buna göre seçim yapacaktır. Okunabilirlik formüllerinin ne olduğu, bu formüllerdeki değişkenler ve Türkçenin yapısal özelliklerinin bütün istisnai kabul edilen varyantlarının tespiti, yapay zekâyâ dayalı bir okunabilirlik formülü için büyük önem taşımaktadır.

Okunabilirlik kavramı alanyazında "anlaşılabilirlik" (Kızılca ve Yılmaz 2008: 72), "okunma zorluk derecesi" (Dale ve Chall 1948: 17; François 2015: 80) veya "okuyucu-metin uyumsuzluğu" (McLaughlin 1969: 640) şekillerinde yer almıştır. Kondru (2006: 7) okunabilirlik formülünü bir metnin okunabilirlik derecesini tahmini olarak veren bir denklem şeklinde tanımlar ve bu tahminin bir kişinin bu metni kavraması beklenen eğitim seviyesinde karşılaşması gerektiğini belirtir. Richards ve Schmidt (2010: 482) okunabilirliği, metnin ne kadar kolay okunabileceğinin ve anlaşılabilirliğinin ölçüsü olarak tanımlamış ve okunabilirlikte ortalama cümle uzunluğu, cümlelerin içerdiği kelime sayısı ve kullanılan dilin ne kadar karmaşık olduğunu yansıtan değişkenler üzerinde durmuştur. Gunning'e göre (1952: 38) okunabilirlik, metnin okunmasını engelleyen şeyin ne olduğunun anlaşılmasını sağlayan pratik bir yoldur. Dubay (2004: 3) okunabilirliğin bazı metinlerin daha kolay okunabilmesini sağladığını ve bu kavramın zaman zaman okunaklılık (legibility) ile karıştırıldığını belirtir.

Bir metnin "kime göre" ya da "nasıl okunabilir" olarak değerlendirilebileceği konusunda Klare (1975: 62) hedeflenen okuyucu kitlesinin "metni çabuk ve açıkça" anlayabildiği derecede okunabilir, olduğunu ifade eder. Bir metin okunabilir olmazsa ne olur, sorusunun cevabını Kondru (2006: 7) çalışmasında okuyucu-metin uyumsuzluğunun okuyucunun metni etkili bir şekilde kullanılamamasına neden olabileceği şeklinde verir. Yazar *Othello* adlı eser üzerinden örnek vererek bu eserlerin seviyelendirilmiş farklı formatlarının olmasının gerektiğini ve eserlere okunabilirliklerine göre seçim atanmasının gerektiğini ifade etmiştir. Collins-Thompson'a göre (2014: 3) pek çok çalışmada metnin zorluğu (text difficulty) ile okuma zorluğunun (reading difficulty) birbirine karıştığını görmek mümkündür.

Yapay Zekâyâ Dayalı Okunabilirlik

Birçok araştırmada okunabilirlik ve anlaşılabilirlik iki farklı kavram olarak ele alınır (Ateşman, 1997; Karatay, Bolat ve Güngör, 2013; Erol, 2014; Baş ve Yıldız, 2015). Buna göre okunabilirlik kelime ve söz dizimi verileriyle ilgili metnin nicel boyutuyla ilgili iken anlaşılabilirlik ise kelimelerin anlamı, mecazi kullanımlar gibi nitel boyutla alakalıdır.

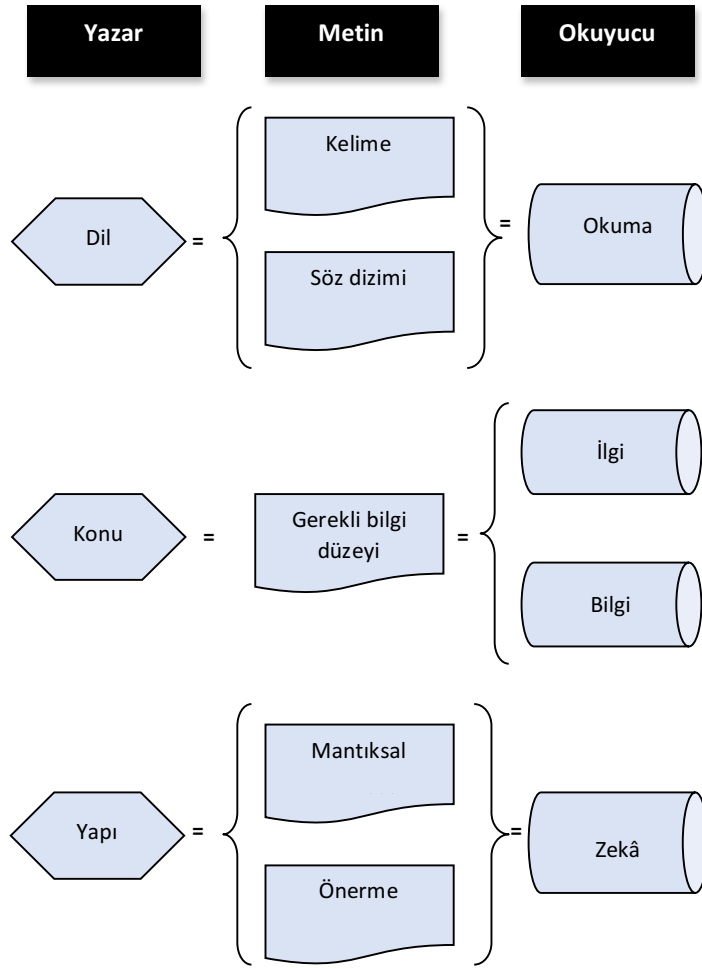
Son dönem çalışmalarda bu iki kesin ayrımı tek bir potada eritme uğraşı öne çıkmıştır. Bu uğraşta yeni değişkenler de eklenerek daha fazla veri unsuru sağlanmıştır. Jones (1997: 106) göre bir metnin anlaşılabilir olması için okunabilir olması gerekir ve okunabilirlik sadece söz dizimsel zorluk değil, okuyucunun bildikleri, ilgileri, okuma düzeyi gibi durumları olarak ele alınmalıdır. Farklı çalışmalarda daha farklı değişkenler işe koşulmuştur. Örneğin Todirascu, François, Gala, Fairon, Ligozat ve Bemhard'e göre (2013: 12, 17) bir metnin iki önemli özelliği olarak tutarlılık ve bağdaşıklık olduğunu belirtip okunabilirliği teorilerini bu iki kavram üzerinden geliştirirler. Yazarlara göre tutarlılık ve bağdaşıklık bir metin okunabilirlik düzeyini tahmin etmede çok önemlidir. Chen ve Meurers (2019), François (2015) ve Vajjala, Meurers, Eitel ve Scheiter'in (2016) çalışmalarında doğal dil işleme ve makine öğrenmenin birleştirildiği modern okunabilirlik çalışmalarının çok kapsamlı olduğunu bireysel öğrenme ihtiyaçlarından, öğrenme görevlerine, öğrenme stillerine ve yabancı dil öğrenme kapsamında dil yeterliliklerine kadar pek çok alanda fayda getirdiğine dikkat çekmişlerdir.

Collins-Thompson'a göre (2014: 6) geleneksel okunabilirlik formülleri bir metnin yüzeysel özelliklerini dikkate alırken bağdaşıklık, söz dizimi belirsizlikleri, mecazlı kullanımlar, deyimler gibi derin yapı özelliklerini dikkate almamaktadır. Yazara göre François'in 2009 yılında oluşturduğu yapay zekâya dayalı formülü (Artificial Intelligence (AI) readability) bu sıkıntıları ortadan kaldırmaktadır. 2012 yılında François ve Fairon bir metnin okunabilirliğine dair iki kavramı işe koşmuşlardır. Bunlar doğal dil işleme (NLP) ve makine öğrenme (ML). Yazarlar buradan hareketle yeni bir kavram ortaya atmışlardır: Yapay Zekâ Okunabilirliği (Artificial Intelligence (AI) Readability) (François 2015: 80).

Yapay Zekâ ile Metin ve Okuyucu Değişkenleri

Geleneksel okunabilirlik çalışmalarında yalnızca metnin özellikleri üzerinde durulmaktadır. Yapay zekâ ile birlikte okuyucu değişkenleri de formüllere dâhil edilmeye başlanmıştır.

Okuyucunun özelliklerinin işe koşulması ve bu özelliklerin metnin özellikleri ile birlikte değerlendirilmesi okunabilirlik formülleri kapsamında yepyeni bir kapıyı aralamıştır. Vajjala, Meurers, Eitel ve Scheiter (2016: 11) ARA adını verdikleri modeli ile okuyucu ve metnin özelliklerini birleştirmeği amaçlamış okumaya ayrılan zaman, okuma yeterlilikleri gibi kavramları işe koşmuşlardır. Collins-Thompson'e göre (2014: 3) bir metnin okunabilirliği, aynı zamanda bir okuyucunun kendine has özelliklerine bağlıdır. Bu özellikler eğitim, sosyal hayatı, ilgi ve tecrübeleri, öğrenme motivasyonu. Bütün bu özellikler bir metnin okunabilirliğinde hayati önem taşır. Yeni nesil okunabilirlik formüllerinde okuyucu faktörünün de devreye girmesiyle okuyucu, metin, yazar bağlamında nelere ihtiyaç olduğu da alanyazında tartışılmaya başlamıştır.



Şekil 1: Okuma Kolaylığı için Gerekli Eşleşmeler

Okunabilirlik çalışmalarından 2009 yılında teorisinin ortaya atılmasıyla birlikte uygulamaya dönük çalışmaları devam etmektedir. Aşağıda ise okunabilirliğe dair 1948 yılından günümüze kadar yapılan uygulamalar yer almaktadır. Türkçe metinler üzerinde bu formüllerin bazen bir tanesi bazen de birden fazlası kullanılmıştır. Aşağıda hem formülleri tanıma hem de hangi formüllerin kullanıldığını görme fırsatı bulacağız.

1.1. Okunabilirlik Formülleri

Bir metnin okunabilirlik tespitine yönelik alanda birden fazla formül geliştirilmiştir. Aşağıdaki tablo okunabilirlik formüllerini tanımak ve hangi formülün Türkçe metinler üzerinde kullanıldığını görmek açısından önem taşımaktadır.

Tablo 1: Okunabilirlik Formülleri ve Formüllerin Kullanıldığı Çalışmalar

Formül	Formülün özelliği	Formüllerin kullanıldığı çalışmalar ²
1. Dale-Chall (1948)	Bir metnin ya da kitabın ilk 100 kelimesi alınır. 100 kelimenin kaç cümleden oluştuğu tespit edilir ve bu kelimeler içindeki bilinmeyen (zor) kelime sayısı hesaplanır. Bilinmeyen	Herhangi bir Türkçe çalışmada kullanılmamıştır.

² Okunabilirlik çalışmasının kullanıldığı 24 tez tespit edilmiştir. Bu tezlerden 18'si ders kitaplarındaki metinler veya çocuk kitapları üzerine yapıldığı için ele alınırken reklam, sağlık vb. alanlarda hazırlanan tezler çalışma dışında kalmıştır. Aynı durum makaleler için de geçerli olup eğitim dışı çalışmalar değerlendirilmeye alınmamıştır.

kelimeler, Dale ve Chall'in belirlediği 3000 kelime (<http://www.rfp-templates.com/research-articles/dale-chall-3000-simple-word-list>) dışındakiler olmak zorundadır. Formülü şöyledir:

Cümle uzunluğu= Kelime sayısı/cümle sayısı

Bilinmeyen kelime oranı ise ilk 100 kelimedeki bilinmeyen kelime sayısının geriye kalan kelime sayısına bölünmesi ve ortaya çıkan sonucun 100 ile çarpılmasıyla elde edilir. Ortaya çıkan puanının ifade ettiği okuma eğitim düzeyi ve yaş aralığı Dale-Chall okunabilirlik düzey tablosuna göre belirlenir.

Okuma Düzey Puanı	Okuma Eğitim Düzeyi	Eğitim Aşaması	Yaş
4.9 ve aşağı	1. düzey 2. düzey 3. düzey 4. düzey	İlköğretim ³	5-10
5.0-5.9	5. düzey 6. düzey		10-12
6.0-6.9	7. düzey 8. düzey	Ortaöğretim	12-14
7.0-7.9	9. düzey 10. düzey		14-16
8.0-8.9	11. düzey 12. düzey		16-18
9.0-9.9	13. düzey 14. düzey	Üniversite	18-22
10+	16. düzey ve üstü		22+

2. Flesch, Okuma Kolaylığı (1948)

Formül bir metnin ne kadar karmaşık olduğunu ölçer. Formüle göre şu aşamalar takip edilir:

-İlk 100 kelime seçilir.

- 100 kelimenin toplam kaç heceden oluştuğu hesaplanır ve hece sayısı kelime sayısına (100) bölünerek metnin ortalama kelime uzunluğu bulunur.

- 100 kelimenin kaç cümleyi oluşturduğunu hesaplanır ve cümle sayısı toplam kelime sayısına (100) bölünerek metnin ortalama cümle uzunluğu bulunur.

- Her iki uzunluk da aşağıdaki formüle yerleştirilerek okuma kolaylığı bulunur.

Okuma Kolaylığı Puanı = $206,835 - (1,015 \times \text{cümle uzunluğu}) - (84,6 \times \text{kelime başına düşen ortalama hece sayısı})$

Puanın ifade ettiği okunabilirlik düzey tablosu şöyledir:

Okuma Kolaylığı Puanı	Okunabilirlik Düzeyi	Metin Örneği	100 Kelimeye Düşen Hece	Ortalama Cümle Uzunluğu u
0-29	Çok Zor	Bilimsel	192+	29+
30-49	Zor	Akademik	167	25
50-59	Oldukça Zor	Nitelikli eserler	155	21

³ Yazarın tablosundaki bu seviyeler ilkokul, ortaokul ve lise/üniversite şeklinde geçmektedir ancak bu durum Türkiye'deki yaş grubuyla uyuşmamaktadır. Bu sebeple makalenin bu kısmında yazarın yaş seviyelendirmesi dikkate alınmıştır.

60-69	Ölçünlü	Seçkin/derleme kitaplar	147	17
70-79	Oldukça Kolay	Kurgu roman	139	14
80-89	Kolay	Çizgi roman	131	11
90-100	Çok Kolay	Karikatürler	123 ya da daha az	8 ya da daha az

3. Gunning, Fog (1952)

Formül dünya çapında çok satan gazete ve dergilerin okunabilirliğinin ölçülmesinde kullanılmaktadır.

İlk100 kelimelik kısım belirlenir ve bu kısımdaki toplam kelime ve cümle sayısı hesaplanır. Formül 3 aşamada tamamlanır:

Birinci aşama: Üç heceli kelime oranı= Üç ve daha fazla heceli kelime sayısı/Geriye kalan kelime sayısı x 100

İkinci aşama: Kelime ortalaması= Kelime sayısı/Cümle sayısı

Üçüncü aşama: Üç heceli kelime oranı + Kelime ortalaması) x 100 = Okunabilirlik Puanı

$$Fog = 0,4x \left(\frac{Kelime\ Sayısı}{Cümle\ Sayısı} + 100x \frac{Çok\ Heceli\ Kelime\ Sayısı}{Kelime\ Sayısı} \right)$$

Formüle göre kısa cümleler daha iyi okunabilirlik seviyesine sahiptir. Fog testine göre 12+ puan üzeri zor metindir. En uygun okunabilirlik puanı 7 ve 8 olarak belirtilmektedir (Dubay 2004: 24).

1. Ateşal (2016)
2. Bezirci ve Yılmaz (2010)
3. Çakmak ve Çil (2014)
4. Günhan (2004)
5. Güven (2010).
6. Kalın (2017)
7. Sönmez (2003)

4. Senter ve Smith, ARI (1967)

ARI formülü ise şu aşamalardan oluşur:

- Birinci aşamada ele alınan metindeki toplam harf sayısı, kelime sayısına bölünür ve 4,71 ile çarpılır.
- İkinci aşamada metindeki toplam kelime sayısı, toplam cümle sayısına bölünür ve 0,5 ile çarpılır.
- Üçüncü aşamada ise ilk iki aşamada elde edilen sonuçlar toplanır ve bu sonuçtan 21,43 çıkarılır. Ortaya çıkan sonuç, metnin okunabilirliğini ifade eder.

$$ARI = 4,71x \left(\frac{Harf\ Sayısı}{Kelime\ Sayısı} \right) + 0,5x \left(\frac{Kelime\ Sayısı}{Cümle\ Sayısı} \right) - 21,43$$

Puan	Yaş
1	5-6
2	6-7
3	7-9
4	9-10
5	10-11
6	11-12
7	12-13
8	13-14
9	14-15
10	15-16
11	16-17
12	17-18
13	18-24
14	24+

1. Bezirci ve Yılmaz (2010)
2. Çakmak ve Çil (2014)

5. McLaughlin, Smog(1969) Smog formülüne göre ilk olarak metnin başından, ortasından ve sonundan 10’ar cümle alınır. Nokta, soru işareti ve ünlemle biten kelime dizileri cümle olarak kabul edilir.
1. Bezirci ve Yılmaz (2010)
2. Kalın (2017)
3. Kalın ve Koçoğlu (2017)

Toplam 30 cümlelik bu kısımdaki +3 heceden oluşan kelimeler sayılır. +3 heceli kelimeler tekrar ediyorsa onlar da sayılmamalıdır.

Toplam +3 heceli kelime sayısının karekökü tahmin edilir. Örneğin 95 çok heceli kelime varsa bu en yakın 100’e yuvarlanır. Karekökü belirlemek için yuvarlayabileceğiniz düşük ve tam sayı seçilir. Örneğin kelime sayısı 110 ise 121 sayısına yuvarlamaktansa 100’e yuvarlamak tercih edilir.

1	4	9	16	25	36	49	64	81	100	121	144	169
---	---	---	----	----	----	----	----	----	-----	-----	-----	-----

Belirlenen sayının karekökü alınır. Çıkan değer, metnin ABD eğitim sistemine göre tanımlanmış olan seviyesidir. Düzeyler 5. sınıftan itibaren başlar. Tablodaki 13-16 arası seviyeler lisans seviyesini gösterir. 17-18 arası ise lisansüstü eğitimi, 18 üstü ise daha yüksek meslekî yeterlilik seviyelerini gösterir.⁴

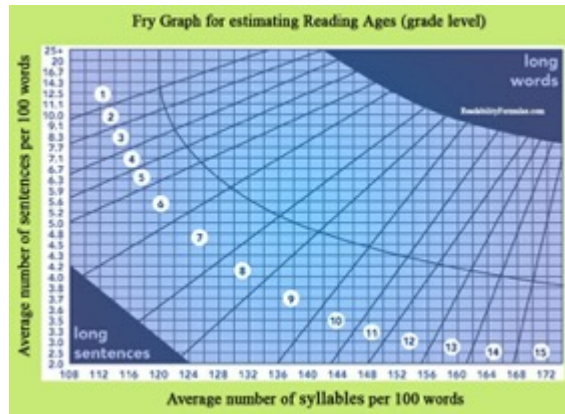
Çok Heceli Kelime Sayısı	Yaklaşık Eğitim Düzeyi
1-6	5
7-12	6
13-20	7
21-30	8
31-42	9
43-56	10
57-72	11
73-90	12
91-110	13
111-132	14
133-156	15
157-182	16
183-210	17
211-240	18

6. Ateşman (1997) Ateşman, Flesch’in (1948) formülünü Türkçe dil yapısına uyarlamıştır. Ateşman’ın (1997) bir metindeki kelime ve cümle uzunluğunu ele alan formülü şu şekildedir:
- O.S.: okunabilirlik sayısı
x1: hece olarak ortalama kelime uzunluğu
x2: kelime olarak ortalama cümle uzunluğu
- $$O.S. = 198,825 - 40,175x_1 - 2,610x_2$$
- Ateşman’ın okunabilirlik düzey tablosu şöyledir:
- | Okuma Metninin Niteliği | Okunabilirlik Sayısı |
|-------------------------|----------------------|
| - Çok kolay | 90-100 |
| - Kolay | 70-89 |
| - Orta güçlükte | 50-69 |
| - Zor | 30-49 |
1. Bağcı ve Ünal (2013)
2. Baki (2019)
3. Baş ve Yıldız (2015)
4. Bezirci ve Yılmaz (2010)
5. Ceran (2015)
6. Çakıroğlu (2015)
7. Çakmak ve Çil (2014)
8. Durukan (2014)
9. Eraslan (2008)
10. Erdem (2011)
20. Özdemir (2016)
21. Polat (2018)
22. Rada (2016)
23. Şimşek (2019)
24. Teke (2016)
25. Temizyürek (2010)
26. Temur (2002)
27. Temur, Sarı ve Orhon (2011).
28. Tosunoğlu ve Özlük (2011)
29. Türkbek (2019)
30. Yılmaz ve Temiz (2014)
31. Ziya (2019)

⁴ <https://readabilityformulas.com/smog-readability-formula.php> sitesinden alınmıştır (Erişim tarihi: 07.02.2020).

- Çok zor	1-29	11. Erol (2014) 12. Hızarcı (2016) 13. İskender (2013) 14. Kalın (2017) 15. Koçağan (2012) 16. Köroğlu ve Zorbaz(2016) 17. Mirzaoğlu ve Akın (2015) 18. Öksüz (2019) 19. Özbek ve Ergül (2018)	32. Zorbaz (2007) 33. Zorlu (2015)
-----------	------	--	---------------------------------------

7. Fry (1977) Fry'ın formülüne göre metnin başından, ortasından ve sonundan alınan yüzer kelimelik toplamda 300 kelimelik bir alandaki toplam cümle, kelime ve hece sayıları aşağıdaki grafikte belirtilen yerlerine konarak metnin sınıf düzeyi belirlenir.
- Birinci aşamada kitap ya da metinden 300 kelimelik bir alan belirlenir.
 - İkinci aşamada cümle sayısı hesaplanır. 100. kelimedeki cümle bitmemişse son cümle içindeki bütün kelimeler cümleye dâhil edilir.
 - Üçüncü aşamada toplam hece sayısı hesaplanır. Hece sayısı ile 300 kelime birbirine bölünür. Heceler, kelimelerdeki ünlü harflerin sayısı kadardır.
 - Son aşamada ortalama cümle uzunluğu grafiğe yerleştirilir ve kesişim noktası belirlenir. Bu nokta yaklaşık eğitim düzeyini ifade eder.



8. Sönmez (2003) Sönmez, metindeki bilinmeyen kelimeler arttıkça anlaşılabilirlik oranının düştüğünü belirtir ve metni nitel özellikleriyle ele alır. Bu yüzden formülünü okunabilirlik değil anlaşılabilirlik olarak ifade eder.
- 100 kelimelik bir kısımdaki toplam cümle sayısı belirlenir. Bilinmeyen kelime sayısı (yabancı kelimeler, terimler,
- 1. Çakmak ve Çil (2014)
 - 2. Elli (2011)
 - 3. Güven (2010)
 - 4. İşeri (2011)
 - 5. Temur, Sarı ve Orhon (2011)

⁵ <https://readabilityformulas.com/graphics/frygraph2lg.jpg> sitesinden alınmıştır (Erişim tarihi: 07.02.2020).

deyimler, mecaz, sembol ve kavramlar) belirlenir. Ardından şu işlemler yapılır:

Anlam oranı= Bilinmeyen kelime sayısı/ cümle sayısı

Kelime oranı= Kelime sayısı/ cümle sayısı

Güçlük oranı= Bilinmeyen kelime sayısı/ kelime sayısı

Anlaşılabilirlik oranı= Anlam oranı / kelime oranı x güçlük oranı

Formülün anlaşılabilirlik düzey tablosu şöyledir:

Anlaşılabilirlik Oranı	Anlaşılabilirlik Düzeyi
1.00-.99	Metin tümüyle anlamsız
.98-.26	Metin anlamsız
.25-.16	Metin bulanık
.15-.09	Metin zor anlaşılır
.08-.04	Metin yardım alınarak anlaşılabilir
.03-.001	Metin anlaşılabilir
.00099-.0001	Metin açık ve anlaşılır
000001-0	Tam iletişim sağlanır.

8. Bezirci ve Yılmaz (2010)

Bezirci ve Yılmaz Türkçe metinlerde 3 ve daha fazla heceli kelimelerin metinde buldukları orana göre değiştiğini belirtir. Yazarlar Ateşman'ın belirttiği Türkçenin ortalama kelime uzunluğunun 2,60 olduğunu doğrular ve tek heceli bağlaçlarının metinlerde çokça geçtiğini ve bu durumun ortalama kelime uzunluğunu da düşürdüğünü ifade ederler. Bu durumda da "Yeni Okunabilirlik Değeri" adlı formülü geliştirmişlerdir. Formüle göre metindeki/kitaptaki her bir cümledeki kelimeler üç, dört, beş, altı+ heceli gibi ayrı ayrı sayılır ve buldukları cümlelere göre ortalamaları hesaplanır.

$$YOD = \sqrt{OKS \times ((H3 \times 0,84) + (H4 \times 1,5) + (H5 \times 3,5) + (H6 \times 26,25))}$$

OKS: Metindeki cümlelerdeki ortalama kelime sayısını ifade eder.

H3: Bir cümledeki ortalama üç heceli kelime sayısı

H4: Bir cümledeki ortalama dört heceli kelime sayısı

H5: Bir cümledeki ortalama beş heceli kelime sayısı

H6: Bir cümledeki ortalama altı veya daha fazla heceli kelime sayısı

Formülde hecelerine göre kelime sayıları ve ortalama kelime sayısı hesaplandıktan sonra ortaya çıkan sonuç Yeni Okunabilirlik Değerini ifade eder. Değerler, 3,03-kolay, 8,30-orta, 18,82-zor şeklinde gruplandırılmıştır. Formüle göre ortaya çıkan sonuçlar aynı zamanda sınıf sonuçlarını da vermektedir; yani bir sonuç 3,10 çıktıysa 3. sınıf, 3,50'nin üstünde çıktıysa 3-4. sınıfa hitap etmektedir.

9. Çetinkaya (2010)

Formülde bir metnin okunabilirliği ortalama cümle ve kelime uzunluğu bulunarak hesaplanır. Metinlerin okunabilirlik düzeyini belirleyen formül denklemi şu şekildedir:

OP= Okunabilirlik Puanı

OTU= Ortalama cümle uzunluğu

OSU= Ortalama kelime uzunluğu

$$OP = 118,823 - 25,987 \times OSU - 0,971 \times OTU$$

Çetinkaya'nın okunabilirlik düzey tablosu şöyledir:

Okunabilirlik Puanı	Okunabilirlik Düzeyi	Eğitim Düzeyi
0-34	Engelli Düzey	10, 11 ve 12. Sınıf
35-50	Eğitsel Okuma	8 ve 9. Sınıf

1. Çakıroğlu (2015)

2. Özbek ve Ergül (2018)

3. Özmen (2019)

1. Bağcı ve Ünal (2013)

2. Baki (2019)

3. Bozlak (2018).

4. Durukan (2014)

5. Güven, Bal ve Halat (2014)

6. Köroğlu ve Zorbaz (2016)

7. Mert (2018)

8. Özbek ve Ergül (2018)

9. Özmen (2019)

10. Polat (2018)

11. Şimşek (2019)

51+	Bağımsız Okuma	5, 6 ve 7. Sınıf	12. Türkben (2019) 13. Ülper, Çetinkaya ve Dikici (2018)
-----	----------------	------------------	---

Araştırma kapsamında İngilizce ve Türkçe çalışmalarda en çok kullanılan dokuz formül tespit edilmiş, formüllerden yalnızca Dale ve Chall herhangi bir Türkçe çalışmada kullanılmamıştır ancak yabancı çalışmalarda çokça kullanıldığı ve alanında ilk olduğu için bu çalışmada kullanılmıştır. Bununla birlikte tablodaki formüllerden Fry'ın okunabilirlik formülü hece sayısına dayandığı için Türkçe uygulamalarında tutarsız sonuçlar vermektedir. Bunun sebebi de Türkçede kelime yapısında her ünlü harf bir heceyi temsil eder. İngilizcede harflerin olmadığı kelimeler de mevcuttur. Bu sebeple Fry formülünde Türkçe metinler çok heceli olarak ortaya çıkmakta ve dolayısıyla Türkçe metinlerin Fry'ın grafiğindeki kesişim noktalarının dışında sonuçları çıkmasından ötürü kullanılmamıştır. Bu formüllerin dışında Powers-Sumner-Kearl, Coleman, Bormuth, Spache, Cloze Test, Raygor, Linsear, Rate Index, Lesbarhets formülleri de yabancı çalışmalarda kullanılmıştır. Okunabilirlik formüllerinin dışında bilgisayar teknolojileri alanında eTma (elektronik Türkçe metin analizi) adlı oluşturulan bir program (Güyer, Temur ve Solmaz, 2009: 758) ve Microsoft Word kelime işlemci programı üzerinden alınan ve Flesch Okuma Kolaylığı ile Flesch-Kincaid Değerine dayanan sonuçların kullanıldığı çalışmalar da bulunmaktadır.

YÖNTEM

Araştırma tarama modeline göre hazırlanmıştır. Tarama modeli, bir durumun şu an olduğu gibi ya da geçmişe dönük ele alınarak betimlenmesini sağlar (Karasar, 2012: 77-79). Tarama modeline yönelik olarak toplanan verilerin organize edilmesinde içerik analizi yöntemi kullanılmıştır. İçerik analizi sadece nicel verilerin değil, nitel verilerin analizinde de kullanılabilir. Bu durumda toplanan veriler okuyucuların daha kolay anlayabilmesi için gruplandırılabilir (Büyüköztürk, Çakmak, Akgün, Karadeniz ve Demirel, 2016: 183). Bu çalışmada alanyazında kullanılan okunabilirlik formülleri, bu formüllerin ele aldığı nicel-nitel değişkenler, Türkçe metinler üzerinde uygulanmış okunabilirlik formülleri ve bu formüllerin 5, 6, 7 ve 8. sınıf ders kitaplarındaki ana metinlere, şiirlere, çocuk kitaplarına uygulanması ve okunabilirlik denemeleri sonrası ortaya çıkan sorunlar temalar altında verilmiştir.

2.1. Araştırmanın İnceleme Nesneleri

Verilerin toplanmasında okunabilirlik formüllerini karşılaştırabilmek için her bir sınıf düzeyindeki ilk üç metin, ilk şiir metni, 6 farklı çocuk kitabı ve araştırma sonuçları desteklemesi açısından akademik bir kitap kullanılmıştır:

Tablo 2: Veri Toplama Kaynakları

Türkçe Ders Kitapları		Çocuk kitapları
Ana metinler	Şiirler	
5. sınıf	1. Hoşça Kalın, Güle Güle 2. Anadolu'da Konukseverlik Gelenekseldir 3. Dumlupınar Savaşı	1. Çıtır Çıtır Felsefe Serisi, Anlaşmak ve Anlaşmamak 2. Pembe İncili Kaftan 3. Pal Sokağı Çocukları 4. Ökkeş Dolmuşçu 5. Başını Vermeyen Şehit 6. Charlie'nin Çikolata Fabrikası
6. sınıf	1. Meşeler 2. Tartışarak... Gerçeğe Doğru 3. Elveda Ağustos Böceği	
7. sınıf	1. Çiçek Dürbünü 2. Ana İşsiz Kalınca 3. Mürefteli Kadınlar ve Emin Astsubay	
8. sınıf	1. İyimserlik ve Kötümserlik Üzerine 2. Kaşığı 3. Bayrağımızın Altında	

Akademik kitap: Albert Einstein, Özel ve genel görelilik kuramı

2.2. Verilerin Toplanması ve Analizi

Araştırmada taranan kaynaklar gerekli tartışmaların ve ortak özelliklerin ortaya konması açısından içerik analizi kullanılmıştır. Araştırmaya konu olan alanyazında geçen okunabilirlik formülleri belirlenmiş, araştırmaya veri sağlayabilecek olan 10 formülden 9'u üzerinden okunabilirlik denemeleri yapılmıştır. Bu denemeler ortaokul Türkçe ders kitaplarında yer alan 12 metin ve ortaokullarda kullanılan 6 farklı çocuk kitabı üzerinde yapılmış, araştırma sonuçlarını desteklemesi açısından da bir akademik kitap kullanılmıştır.

Yapılan içerik analizinde araştırma kapsamında belirlenen okunabilirlik formüllerinde kullanılan nitel ve nicel değişkenlerin formüllerin ölçme sonuçlarını nasıl ele aldıkları ve hangi kitleye hitap ettikleri belirlenmiştir. Bundan sonraki aşamada ise her metin/kitap formüllerin ifade ettiği aşamalara göre denenmiştir.

Araştırma süreci şu şekilde yapılandırılmıştır:

Şekil 1: Araştırma süreci

Okunabilirlik formüllerinin belirlenmesi	<ul style="list-style-type: none"> Alanyazındaki okunabilirlik formüllerinin tespit edilmesi. Türkçe çalışmalarda kullanılan formüllerin ayrı bir havuza alınması. Eğitim çalışmalarında kullanılan formüllerin belirlenerek araştırmada kullanılması.
Veri toplama kaynaklarının belirlenmesi ve uygulamaların yapılması	<ul style="list-style-type: none"> Araştırmada kullanılacak formüller için ortaokul Türkçe ders kitaplarındaki (2019-2020) ilk üç metin ile ortaokullarda kullanılan 6 çocuk kitabının belirlenmesi. Belirlenen formüllerin her bir veri toplama kaynağına uygulanması ve sonuçların tablollaştırılması.
Yapılan denemeler sırasında ortaya çıkan problemlerin belirlenmesi	<ul style="list-style-type: none"> Okunabilirlik denemeleri yapılırken bir değişkenin nasıl sayılacağı ya da hangi değişken grubuna dâhil edileceği vb. sorunların notlanması. Belirlenen her bir sorunun ortak temalar altında birleştirilmesi ve araştırmacıların bu sorunlarla ilgili yaptıkları açıklamalara yer verilmesi.
Problemler doğrultusunda değişkenlere göre farklı sinamaların yapılması.	<ul style="list-style-type: none"> Sayımının nasıl yapılacağı bilinmeyen değişkenlere yönelik her iki duruma göre formüllerin uygulanması.

BULGULAR

Bulgular iki bölüm altında işlenmektedir. Birinci bölümde araştırma kapsamında belirlenen okunabilirlik formülleri değişkenleri ve sonuçları bakımından karşılaştırmalı olarak incelenmiş, formüller metinlere ve kitaplara uygulanmıştır. İkinci bölümde ise formüller uygulanırken yaşanan problemler örnekleriyle ve formüllerdeki yönergeler üzerinden farklı durumlar sınanarak verilmiştir.

3.1. Okunabilirlik Formüllerinde Kullanılan Değişkenler

Formüllerinin metinlerin okunabilirliği belirlemek için ele aldıkları nicel/nitel değişkenleri aşağıdaki tabloda gösterilmiştir.

Tablo 3: Formüllerde ele alınan nicel ve nitel değişkenler

Formüller	Nicel değişkenler	Nitel değişkenler

	<i>Kelime sayısı</i>	<i>Cümle sayısı</i>	<i>Hece sayısı</i>	<i>Harf sayısı</i>	<i>Çok heceli kelime sayısı</i>	<i>Bilinmeyen / zor kelime sayısı</i>
1. Dale ve Chall (1948)	+	+				+
2. Flesch, Okuma Kolaylığı (1948)	+	+	+			
3. Fog (1952)	+	+			+	
4. Senter ve Smith, ARI (1967)	+	+		+		
5. Smog (1969)		+			+	
6. Ateşman (1997)	+	+	+			
7. Fry (1977)	+	+	+			
8. Sönmez (2003)	+	+				+
9. Bezirci ve Yılmaz (2010)	+	+	+			
10. Çetinkaya (2010)	+	+	+			

İlk okunabilirlik çalışmalarından biri olan Flesch'in (1948) kelime ve cümle uzunluğu değişkenlerine dayanmaktadır. Sonraki formüllerin neredeyse tamamında bu şekilde nicel değişkenler kullanılmış olup nitel değişkenler yalnızca Dale ve Chall (1948) ile Sönmez'in (2003) formüllerinde yer almıştır.

3.1.2. Sonuçları bakımından okunabilirlik formülleri

Okunabilirlik formülleriyle atanan düzeyler, metnin kime hitap ettiğine yönelik bilgi verir. Formüllerin düzey bilgileri şöyledir:

Tablo 4: Formüllerin Okunabilirlik Düzey Sonuçları

Okunabilirlik formülleri	Yaş	Sınıf	Metin zorluğu/seviyesi
1. Dale ve Chall(1948)	+	+	
2. Flesch (1948)			+
3. Fog (1952)			+
4. Senter ve Smith, ARI (1967)	+		
5. Smog (1969)		+	
6. Ateşman (1997)			+
7. Fry (1977)		+	
8. Sönmez (2003)			+
9. Bezirci ve Yılmaz (2010)		+	+
10. Çetinkaya (2010)		+	

Tablo 4'te de görüldüğü gibi formüller ya sınıf ya da metin zorluğu üzerinden düzeylerini ifade etmişlerdir. Çok az formül ise yaş bilgisini vermiştir.

3.1.3. Okunabilirlik Formüllerinin Uygulanması

Bu aşamada formüller Türkçe ders kitaplarındaki metinlere ve çocuk kitaplarına uygulanmıştır.

3.1.3.1. Formüllerin Türkçe Ders Kitaplarındaki Metinlere Uygulanması

Bu kısımda formüllerin aynı metinde benzer sonuçlar verip vermeyeceği test edilmiştir. Bu amaçla okunabilirlik formülleri 5, 6, 7 ve 8. sınıf Türkçe ders kitaplarındaki ilk üç ana metin üzerinden işletilmiştir.

Tablo 5: Formüllere göre metinlerin okunabilirlik düzeyleri

Sınıf	Metnin Başlığı	Dale ve Chall	Flesch	Fog	ARI	Smog	Ateşman	Sönmez	Bezirci ve Yılmaz	Çetinkaya
5	1. Hoşça Kalın, Güle Güle	OD ⁶ : 5-6. düzey Eğitim aşaması: İlköğretim Yaş: 10-12	1,8 Metin çok zor	9,5 metin anlaşılabilir	11,98 16-17 yaş	15 eğitim seviyesi	71,44 kolay	AO ⁷ : 0.0485 Metin yardım alınarak anlaşılabilir	7,32 7. sınıf	OD: Eğitsel okuma OED ⁸ : 8 ve 9. sınıf
	2. Anadolu'da Konukseverlik Gelenekseldir	OD: 16. düzey Eğitim aşaması: Üniversite Yaş: 22+	-41,92 Metin çok zor	9,2 metin anlaşılır	16,52 18-22 yaş	17 eğitim seviyesi	54,52 orta güçlükte	AO: 0.01037 Metin zor anlaşılır.	8,19 8-9. sınıf	OD: Engelli okuma OED: 10, 11 ve 12. sınıf
	3. Dumlupınar Savaşı	OD: 16. düzey ve üstü Eğitim aşaması: Üniversite Yaş: 22+	-29,97 Metin çok zor	6,8 metin kolay anlaşılır	14,48 18-22 yaş	14 eğitim seviyesi	65,74 orta güçlükte	AO: 0.08160 Metin yardım alınarak anlaşılabilir	6,22 6. sınıf	OD: Eğitsel okuma OED: 8 ve 9. sınıf
6	1. Meşeler	OD: 7-8. düzey Eğitim aşaması: Ortaöğretim Yaş: 12-14	-24,95 Metin çok zor	8,2 metin kolay anlaşılır	12,74 17-18 yaş	14 eğitim seviyesi	69,27 orta güçlükte	AO: 0.02979 Metin anlaşılabilir	6,40 6. sınıf	OD: Eğitsel okuma OED: 8 ve 9. sınıf
	2. Tartışarak... Gerçeğe Doğru	OD: 1, 2, 3 ve 4. düzey Eğitim aşaması: İlköğretim Yaş: 5-10	-34,99 metin çok zor	12,3 metin zor anlaşılır	14,01 18-22 yaş	14 eğitim seviyesi	66,34 orta güçlükte	AO: 0.03 Metin anlaşılabilir	11,60 10-11. sınıf	OD: Eğitsel okuma OED: 8 ve 9. sınıf
	3. Elveda Ağustos Böceği	OD: 7-8. düzey Eğitim aşaması: Ortaöğretim Yaş: 12-14	-4,70 metin çok zor	7,6 metin kolay anlaşılır	17,40 18-22 yaş	14 eğitim seviyesi	80,57 kolay	AO: 0.33 Metin anlamsız.	8,9 8-9. sınıf	OD: Eğitsel okuma OED: 8 ve 9. sınıf

⁶ Okunabilirlik Düzeyi

⁷ Anlaşılabilirlik Oranı

⁸ Okuma Eğitim Düzeyi

7	1. Çiçek Dürbünü	OD: 1, 2, 3 ve 4. düzey Eğitim aşaması: İlköğretim Yaş: 5-10	-40,94 metin çok zor	11,1 metin anlaşılabilir	14,40 18-22 yaş	14 eğitim seviyesi	65,35 orta güçlükte	AO: 0,01 Metin anlaşılabilir	12,28 12. sınıf	OD: Eğitsel OED: 8 ve sınıf
	2. Ana İşsiz Kalınca	OD: 1, 2, 3 ve 4. düzey Eğitim aşaması: İlköğretim Yaş: 5-10	-19,68 metin çok zor	7,8 metin kolay anlaşılır	10,16 15-16 yaş	11 eğitim seviyesi	77,49 kolay	AO: 0,01 Metin anlaşılabilir	9,5 9. sınıf	OD: Eğitsel okuma OED: 8 ve 9. sınıf
	3. Mürefteli Kadınlar ve Emin Astsubay	OD: 11-12 düzey Eğitim aşaması: Ortaöğretim Yaş: 16-18	-26,50 metin çok zor	10,13 metin anlaşılabilir	11,20 16-17 yaş	14 eğitim seviyesi	74,19 kolay	AO: 0,06 Metin yardım alınarak anlaşılabilir	9,02 9. sınıf	OD: Eğitsel OED: 8 ve 9. sınıf
8	1. İyimsizlik ve Kötümsizlik Üzerine	OD: 9-10. düzey Ortaöğretim Yaş: 14-16	-57,76 metin çok zor	15,15 metin zor anlaşılır	18,46 18-22 yaş	18 eğitim seviyesi	48,95 zor	AO: 0,04 Metin yardım alınarak anlaşılabilir	10,06 10. sınıf	OD: Engelli düzey OED: 10, 11 ve 12. sınıf
	2. Kaşığı	OD: 16. düzey ve üstü Eğitim aşaması: Üniversite Yaş: 22+	-19,86 Metin çok zor	15,15 metin zor anlaşılır	9,60 14-15 yaş	13 eğitim seviyesi	80,63 kolay	AO: 0,09 Metin zor anlaşılır.	7,002 7. sınıf	OD: Eğitsel okuma OED: 8 ve 9. sınıf
	3. Bayrağımızın Altında	OD: 5-6. düzey Eğitim aşaması: İlköğretim Yaş: 10-12	-15,47 Metin çok zor	12,52 metin kolay anlaşılır	11,76 16-17 yaş	14 eğitim seviyesi	72,39 kolay	AO: 0,02 Metin anlaşılabilir	12,03 12. sınıf	OD: Eğitsel okuma OED: 8 ve 9. sınıf

Tablo 5'te de görüldüğü üzere aynı metin hakkında okunabilirlik formülleri farklı sonuçlar vermiştir.

3.1.3.2. Formüllerin Çocuk Kitaplarına Uygulanması

Bu kısımda araştırma kapsamında belirlenen 6 farklı kitaba formüllerin her biri uygulanmıştır:

Tablo 6: Formüllere göre çocuk kitaplarının okunabilirlik düzeyleri

Kitabın Adı	Dale ve Chall	Flesch	Fog	ARI	Smog	Ateşman	Sönmez	Bezirci ve Yılmaz	Çetinkaya
1. Çıtır Çıtır Felsefe Serisi, Anlaşmak ve Anlaşmama k	5.2 OD: 5-6. düzey Eğitim aşaması: İlköğretim Yaş: 10-12	16.10 Metin çok zor	8,7 Metin zor anlaşılır.	6,8 5-6 yaş	6,7 11 eğitim seviyesi	84,53 Kolay	AO: 0,39 Metin açık ve anlaşılır.	3,48 3. sınıf	48,48 OD: Eğitsel okuma OED: 8 ve 9. sınıf
2. Pembe İncili Kaftan	8.69 OD: 11 ve 12. düzey Eğitim aşaması: İlköğretim	22,41 Metin çok zor	11+ Metin zor anlaşılır.	11,98 16-17 yaş	15 eğitim seviyesi	57,786 Orta Güçlükte	AO: 1,01 Metin anlamsız.	8,8 8-9. sınıf	34,11 OD: Eğitsel okuma OED: 8 ve 9. sınıf

	Yaş: 16-18									
3. Pal Sokağı Çocukları	6,38	38,98	11+	11,98	13 eğitim seviyesi	58,74	AO: 0,01	8,9	32,6	
	OD: 7 ve 8. düzey	Metin zor	Metin zor anlaşılır.	16-17 yaş		Orta Güçlükte	Metin zor anlaşılır.	8-9. sınıf	OD: Eğitsel	OED: 8 ve 9. sınıf
	Eğitim aşaması: İlköğretim									
	Yaş: 12-14									
4. Ökkeş Dolmuşçu	7,69	37,79	34,28	11,98	13 eğitim seviyesi	60,18	AO: 0,39	8,1	33,45	
	OD: 9 ve 10. düzey	Metin zor	Metin zor anlaşılır.	16-17 yaş		Orta Güçlükte	Metin açık ve anlaşılır.	8. sınıf	OD: Eğitsel	OED: 8 ve 9. sınıf
	Eğitim aşaması: İlköğretim									
	Yaş: 14-16									
5. Başını Vermeyen Şehit	7,78	36,43	11+	11,98	14 eğitim seviyesi	60,55	AO: 0,35	8,74	37,35	
	OD: 9 ve 10. düzey	Metin zor	Metin zor anlaşılır.	16-17 yaş		Orta Güçlükte	Metin açık ve anlaşılır.	8-9. sınıf	OD: Eğitsel	OED: 8 ve 9. sınıf
	Eğitim aşaması: İlköğretim									
	Yaş: 14-16									
6. Charlie'nin Çikolata Fabrikası	4,69	35,96	9,9	6,8	6,5	98,07 Çok Kolay	AO: 0,04	7,53	40,65	
	OD: 1, 2, 3 ve 4. düzey	Metin zor	Metin zor anlaşılır.	5-6 yaş	11 eğitim seviyesi		Metin açık ve anlaşılır.	7-8. sınıf	OD: Eğitsel	OED: 8 ve 9. sınıf
	Eğitim aşaması: İlköğretim									
	Yaş: 5-10									

Tablo 6'da görüldüğü üzere düzeyler birbirinden farklı çıkmaktadır. Bazı formüllerde cümleleri kısa olan kitaplarda, kitaplar ortaokul için uygun görülse/kullanılsa da okunabilirlik yaşı 5-6 olarak çıkmaktadır.

3.2. Formüllerdeki Değişkenler ve Türkçenin Yapısal Özellikleri

Türkçenin ifade kabiliyeti aynı zamanda yazım çeşitliliğini de beraberinde getirir. Yazım çeşitliliği okunabilirlik formüllerinin doğru işletilebilmesinde önem taşımaktadır. Bu aşamada okunabilirlik formüllerini işletirken farklı, özel veya istisnai olarak görülen çeşitlilikler tespit edilmiştir. Türkçenin yapısal özellikleri formüllerdeki değişkenler üzerinden nicel ve nitel olmak üzere iki alt başlık altında yer almaktadır.

3.2.1. Nicel Kısımla İlgili Problemler

Nicel kısımla alakalı problemler Türkçenin yapısal özellikleri üzerinden 11 ana temada yer almaktadır:

1. Cümle tanımı

Okunabilirlik formüllerinde cümle sayısı en önemli değişken olarak yer almaktadır. Bu cümle tiplerinin nasıl kabul edileceği varsayımı çıkan sonuca önemli etki edecektir. Türkçede cümle değişkenin çok

farklı görünümleri vardır. Bu noktada bilgisayara tanımlanacak cümle tipleri çeşitlilik arz etmektedir. Cümle tanımını etkileyen değişkenler şunlardır:

1.1. İç içe birleşik cümlelerin sayımı

Problem: Çift tırnak içinde birden fazla cümle varsa bu cümleler, cümle sayısına dâhil edilecek mi?

Okunabilirlik formüllerinde çift tırnak içindeki cümlelerin nasıl ele alınacağı kısmı farklı durumlar belirtilerek açıklanmamıştır. Çetinkaya (2010) formül çalışmasında çift tırnak içindeki cümlelerin de ayrı bir cümle olarak alınması gerektiği belirtilmiştir. Bu durumda ise şu problemler yaşanmaktadır:

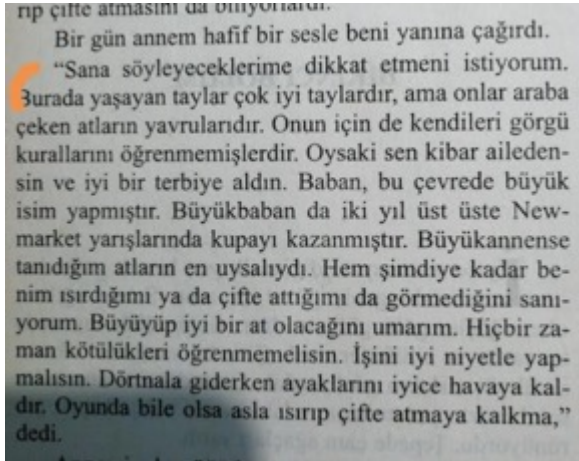
Örneğin 6. sınıf ders kitabındaki *Elveda Ağustos Böceği* metninde geçen aşağıdaki cümle tek bir cümle olarak alındığında okunabilirliği yüksek, dört farklı cümle olarak alındığında okunabilirliği düşük olarak iki uç noktada sonuçlar çıkmaktadır:

İrân ve ayın gülmüşeği yaz gecelerinde, sizin şarkılarınızı dinlemeliydik biz. Bazen kocaman ormanların derinliklerinden, bazen de evlerimizin bahçesindeki yaşlı bir iğde ağacının dalları arasından seslenmeliydiniz:

“Allah’ım dünyanın ne güzel! Dünyan ne güzel Allah’ım! Sen ne güzelsin!”

Bazımız duyacak, bazımız duymayacak;

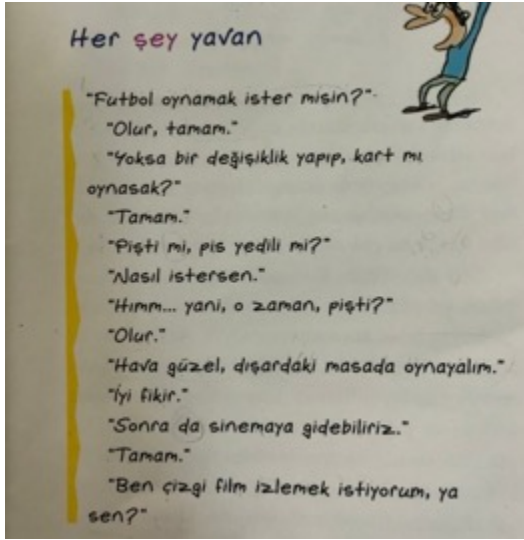
Bu durumun en belirgin örneği ise çocuk kitaplarında görülmüştür. Burada ise çift tırnak içerisinde birçok cümle bulunmaktadır. Peki, bu noktada çift tırnak nasıl tanımlanacaktır?



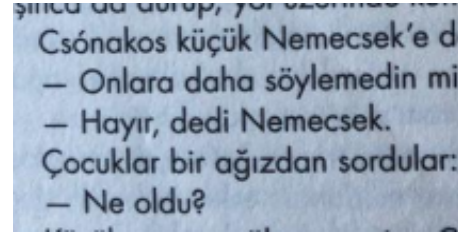
np çifte atmasını da biliyorlar.
Bir gün annem hafif bir sesle beni yanına çağırdı.
“Sana söyleyeceklerime dikkat etmeni istiyorum.
Burada yaşayan taylorlar çok iyi taylorlardır, ama onlar araba çeken atların yavrularıdır. Onun için de kendileri görgü kurallarını öğrenmemişlerdir. Oysaki sen kibar ailedensin ve iyi bir terbiye aldın. Baban, bu çevrede büyük isim yapmıştır. Büyükbaban da iki yıl üst üste Newmarket yarışlarında kupayı kazanmıştır. Büyükannense tanıdığım atların en uysalıydı. Hem şimdiye kadar benim ısırığımı ya da çifte attığımı da görmediğini sanıyorum. Büyüyüp iyi bir at olacağını umarım. Hiçbir zaman kötülükleri öğrenmemelisin. İşini iyi niyetle yapmalısın. Dörtlüye giderken ayaklarını iyice havaya kaldır. Oyunda bile olsa asla ısırıp çifte atmaya kalkma,” dedi.

Bu cümlede çift tırnak ile başlayan kısımdan tırnağın kapanışına kadar bir cümle olarak ele alındığında sondaki “dedi” kısmına kadar 12 cümle bulunmaktadır. Çift tırnak içindeki ifadelerin tek ya da birden fazla cümle olarak alınması arada ciddi nicel farklılıkları ortaya çıkaracaktır (*Charlie'nin çikolata fabrikası*, 13. s.).

Farklı bir durum ise *Çıtır Çıtır Felsefe: Anlaşmak ve Anlaşmamak* (3. s.) adlı kitapta görülmüştür. Kitaptaki cümleler çok kısa olduğu için okunabilirliği seviyesi 5-6 yaş olarak çıkmıştır. Bu durum cümlelerin kısa olmasından yer yer de tek bir kelimedenden oluşmasından dolayıdır. Buna benzer bir durum *Pal Sokağı Çocukları* (25. s.) adlı kitapta da görülmektedir. Burada da cümleler konuşma çizgisi ile verilmiştir. Örneğin “Çocuklar bir ağızdan sordular: -Ne oldu?” ifadesi iki cümle olarak mı alınacak tek bir cümle olarak mı kısmı bir sorun oluşturmaktadır.

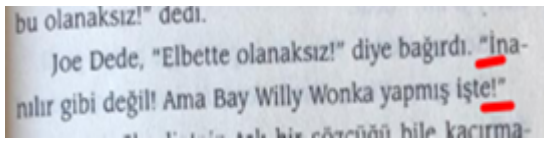


Çıtır Çıtır Felsefe serisi: Anlaşmak ve Anlaşamamak



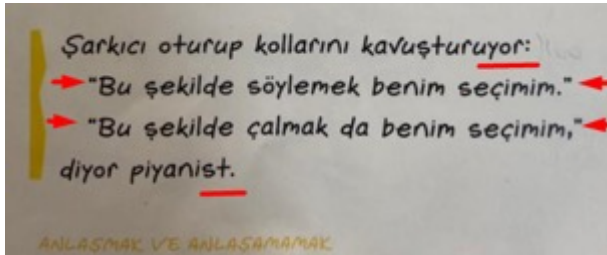
Pal Sokağı Çocukları

Çift tırnak içinde cümlelerin sayımının yanı sıra çift tırnak dışındaki kısımların cümle sayısına nasıl dâhil edileceği kısmı da bilinmemektedir. Örneğin *Charlie'nin Çikolata Fabrikası*(6. s.) kitabındaki şu cümlelere bakalım:



Yandaki resimde yer alan ilk cümlelerin içindeki aktarma cümle dışta kalan "Joe Dede" ve "diye bağırdı" ifadesi bir cümle olarak mı kabul edilecek? Ayrıca ikinci kısım ise sadece tırnak içinde olan iki cümleden oluşmaktadır. Bu durumda bu kısım iki, üç ya da dört cümle olarak alınması sorunu doğmaktadır.

Yine *Çıtır Çıtır Felsefe: Anlaşmak ve Anlaşamamak* (5. s.) kitabında ilk ifade iki nokta ile bitiyor ve kitabın bağlamından anlaşıldığı kadarıyla ilk çift tırnak içi cümle, bu kısma ait. İkinci çift tırnak içi cümle ise piyaniste aittir:



Yandaki görselde yer alan ifadelerin Türkçe yazım kuralları içerisinde tek cümle olarak alınması gerekmektedir. Çift tırnak sonrası olası seçenekler yapay zekâya ayrıntılı tasvir edilmelidir.

Benzer durum, çift tırnak içindeki cümlelerin bir cümle bitiricisiyle bitmemesi kısmında da rastlanmıştır.

Yüzlerce ve yüzlerce şey konusunda insanlar aynı fikirde olmaya çalışmazlar ve her şey yolunda gider. Birini üzgün hissetmeye, terlemeye ya da soyut sanatı sevmeye ikna etmeye çalışmak saçmadır.
"Herkesin bisleri kendine, herkesin duygularını kendine, herkesin zevkleri kendine," demekte haklıyız.

Yandaki resimde görüldüğü gibi virgülün bir cümle bitirici olmadığı bilindiğinde tırnak içi ifadenin de bir cümle olarak değerlendirilemeyeceği varsayılmaktadır. Devam eden kelimelerin ayrı bir cümle olarak alınıp alınmayacağı bir sorundur(6. s.).

1.1.1. Çift tırnak içindeki kelimelerin cümlelerden ayrılması

Problem: Çift tırnak içindekiler cümle mi yoksa kelime mi?

Çift tırnak işareti bilgisayara tanıtılırken içindekinin cümle mi yoksa kelime mi olduğu bilgisi cümle sayısını tanımlamada önemlidir:

Diğerleri ne konuştuklarını anlamak için soze karıştılar.
— Dün bize Müze'de "Einstead" yaptılar.
"Einstead" Almanca bir sözcüktü, en kuvvetlinin hakkını simgeliyordu. Budapeşte'li çocuklar için çok özel bir anlamı vardı. "Einstead" diye bağırıldığında,

Yandaki resimde de görüldüğü üzere çift tırnak sadece cümlelerde değil, kelimelerde de kullanılabilir (7. s.).

1.2. Sıralı cümlelerin sayımı

Problem: Virgülle ya da noktalı virgülle ayrılan sıralı cümleler tek mi yoksa birden fazla cümle mi sayılacak?

Makinelerle dolu bir dünya

Fırat, çalar saatine vurup mutsuzluk çarılarını susturuyor. Bedeninin yarısı yataktan düşüyor ve müziği sonuna kadar açıyor. Duşa giriyor, sütü buzdolabından çıkarıp ısıtıyor, alelacele tıraş oluyor, 2 dilim ekmeği kızartıyor, Facebook'a giriyor, kâsesini bulaşık makinesine koyuyor, Selin'e bir SMS atıyor, kurutma makinesinde kot pantolonunu arıyor, hızlı dişlerini fırçalıyor, tuvalete giriyor ve onu okula götürecek otobüsü yakalamak için kulağında kulaklıkla konuşmaya başlıyor.

Türkçede öznesi ya da yüklemi ortak olan cümleler tek bir cümle olabilir. Yandaki resimde birden fazla cümlelerin virgülle bağlandığı dil bilgisi olarak tek bir cümle olarak kabul edilen örneğe rastlanılmıştır. Bu cümleler tek tek alındığında okunabilirlik seviyesi çok yüksek çıkmasına rağmen tek bir cümle olarak alındığında ise 44 kelimelik uzun ve okunabilirliği çok düşük bir cümle ortaya çıkmaktadır.

1.3. Bağlı cümlelerin sayımı

Problem: Bağlaçlarla ayrılan bağlı cümleler tek mi yoksa birden fazla cümle mi sayılacak? Bağlaçlarla bağlı cümleler bazen tek bir cümle olarak alınırken bazen iki cümle olarak alınmaktadır. Bağlaçların bir cümlede varlığı bilginin doğru anlaşılmasında önemlidir. Bağlaç bilgisi atlandığında anlam farklılaşabilir. Türkçenin yapısal özellikleri açısından bakıldığında aşağıdaki örnekler önem arz etmektedir. *Pal Sokağı Çocukları* (23-24. s.) adlı kitaptaki aşağıdaki ilk cümle bitmiş, ikinci cümle bir bağlaçla bağlanmıştır ancak yine aynı kitapta başka bir cümle şu şekildedir:

yemeyi göze aldığı anlamını tanımlanıyor. Ama aynı zamanda, olağanüstü, kuşatma, birine karşı güç kullanma, birlik yapma gibi birçok kavramın da kısaca

olmasıydı. Nemecek yukarı baktı. Kimseyi gözü kaleyi oluşturan odunların arasında birini hissetti. Odunlara tutunarak yukarı tırman

Yukarıdaki iki cümlede de “ama” bağlacı iki ayrı cümleyi bağlamaktadır. Yapay zekâya bu bilgi tanıtılırken virgül veya bağlaç cümle bitirici olarak tanımlanmayacağı hâlde tek cümlenin anlam yoğunluğu ile iki cümlenin anlam yoğunluğu aynı olamaz. Bu noktada nicel gibi görünen bağlı cümle bir kısmıyla da formüllerin nitel kısımlarıyla alakalıdır.

1.4. Büyük harf kullanımı

Büyük harf kullanımı, cümle değişkeninin de bir belirleyicisidir. Cümle, yapay zekâya tanıtılırken “Büyük harfle başlar, cümle bitirici ile bitirilir.” şeklindedir. Peki, büyük harf kullanımı yapay zekâya nasıl tanıtılmalıdır?

“Çünkü öyle.” Çat!
İnsanın suratına kapı gibi çarpan bir yanıt.

“Boş ver, herkesin fikri kendine.”

“Herkesin fikri kendine.” Çat!

Yandaki resimde bulunan iki cümlede çift tırnak içinde cümle bulunmaktadır ancak bu cümlelerden sonra “Çat!” ünlemi gelmektedir. Bu cümlelerin ayrı ya da birlikte alınması problemi için noktalama işaretlerinin doğru şekilde kullanılması gerekmektedir(6. s.).

Buradaki problem yapay zekânın her çift tırnak sonrası gelen kelimenin devam cümlesi mi yoksa buradaki gibi yeni bir cümle mi olduğu ayrımını yakalamasıdır.

2. Sayı, tarih ifadeleri ve sembol/simgelerin sayımı

Problem: Sayılar tek mi yoksa birden fazla kelime mi? Sayıların harfle ya da rakamla yazımı arasındaki fark nasıl sayılmalıdır? Metinlerde yer alan sembollerin hece sayımı okunuşa göre mi yoksa sembol üzerinden mi yapılacaktır?

Sayıların okunduğu gibi alınması hece sayısını etkileyerek ortalama kelime uzunluğunu artırmakla birlikte bu ifadelerin tek kelime olarak sayılması da ortalama cümle uzunluğunu kısaltarak sonuç olarak okunabilirlik seviyesi yükseltmektedir. İkinci sorun sayıların rakam yerine yazıyla yazılmasıdır. Sayıların yazıyla yazıldığında kaç kelime veya hece alınacağı net olarak ifade edilmelidir: *yüz dört* veya *104*. Bu iki kullanımın aynı bilgi aktarılacak şekilde ilk iki kelime ikincisi tek kelime olarak kabul edilmektedir. Örneğin;5. sınıf ders kitabındaki *Dumlupınar Savaşı* adlı metninde şu cümle geçmektedir: “Mustafa Kemal, 23 Ağustos 1922 günü gizlice Konya’ya gitti.”. Bu cümle9 kelime olarak alındığında ortaokul ders kitapları için uygun cümle uzunluğunda olmasına rağmen okunduğu şekilde alındığında 14 kelime olmaktadır. Bu durumda Flesch ve Çetinkaya (2010) formüllerine göre okunabilirlik düşmekte, Fry’ın formülüne göre okunabilirliği yükseltmektedir. Bununla birlikte metin içlerinde rastlanan $\frac{3}{4}$ (üç bölü dört), %100 (yüzde yüz) benzeri kullanımların tek bir kelime olarak mı yoksa okunduğu şekilde mi sayılacaktır? Konuyla alakalı olarak Flesch (1948: 228) sembol (Örneğin, \$ “dolar”) okunduğu gibi sayılması gerektiğini ancak bir paragrafta birden fazla çok basamaklı sayılar varsa bunların okunuşunun toplam hece sayısına dâhil edilmemesi gerektiğini belirtmiştir.

3. Kısaltmaların sayımı

Problem: Kısaltmalar tek bir kelime mi yoksa kısaltmayı oluşturan her bir kelime kadar mı sayılacaktır? Kısaltmalar kısaltıldıkları şekilde tek bir kelime olarak sayılırsa hece sayısı da harf sayısı kadar mı sayılacaktır?

TDK’ye göre büyük harf yazılan kısaltmalar (TBMM), kısaltıldıkları şekilde (TeBeMeMe) alınmalıyken, küçük harfle yazılan kısaltmalar (av.) okunduğu gibi (avukat) alınmalıdır. Aşağıdaki üç resimde

görüldüğü üzere kitaplarda farklı türde kısaltmalara rastlanmıştır. Bunlardan ilki MP3 (MEPE3), ikincisi ODTÜ (Orta Doğu Teknik Üniversitesi), üçüncüsü ise OKS'dir (Ortaöğretim Kurumları Seçme ve Yerleştirme Sınavı):

lar ve yararsızlar... Bilgisayar gibi harika bir şey yararsız mı oluyor bu durumda? Sırf biz çok zaman veriyoruz diye cep telefonları, MP3-çalarlar, ne bileyim işte, sürekli kullandığımız üstün teknolojiler yararsız mı? Hayır, hep-

ederim. Ablam ODTÜ'yü kazanıp da öğrenci yurduna yerleşince nasıl rahatlamıştı. "Annemin sesi buradan duyulmuyor, Selim. İnanabiliyor musun?" demişti bana daha

Sekizinci sınıftayım. OKS yüzünden halimiz duman! Anadolu Lisesi'ne kapağı atamazsam yandım. Babam

Kısaltmalar okunduğu şekilde alındığında cümle ve kelime uzunluğu artacaktır. Konuyla ilgili olarak formüllerde kısaltmaların sadece harf olarak sayılması ya da okunduğu gibi alınması kısımlarına yönelik bir yönerge belirlenmelidir. Bununla birlikte kısaltıldıkları şekilde ele alındıklarında da hece sayısının okunuşuna göre mi sayılıp sayılmayacağı bilinmemektedir. Bu durumda formüllerin büyük ve küçük harfli kısaltmaları dikkate alarak bir yönerge belirlemeleri ve yapay zekâda da bunu belirtmeleri gerekmektedir.

4. Edat-bağlaçların sayımı

Problem: Edat ve bağlaçlar ayrı bir kelime olarak mı alınacak? Bitişik yazılan ve iki kelimededen oluşan edat/bağlaçların sayımı nasıl ele alınacak?

Birinci sorun, Türkçe eklemeli bir dil olduğu için bazı edat ve bağlaçlar ayrı ya da bitişik yazılabilmektedir. Örneğin "ile" kelimesi bir edat olarak kullanıldığında kelimeye eklenmektedir. "ile" bağlacının yanı sıra özellikle cümle sonlarında rastlanılan "mı?", "mıdır?" veya "mısınız?" edatlarının da ayrı bir kelime olarak alınıp alınmayacağı bilinmemektedir.

Okunabilirlik formüllerinin hepsinde karakterler arasındaki her bir boşluk bir kelime olarak alınmaktadır. Bağlaçlarda problem bu noktada ortaya çıkmaktadır. İki kelimededen oluşan bağlaç ya da edatların tek mi yoksa iki kelime mi olacağıdır: *aksi takdirde* veya *ya da*.

5. Çok heceli kelime ve hece ortalaması

Problem: Harf sayısı az olup hece sayısı fazla olan kelimelerin sayımı nasıl yapılacaktır?

Smog (1969) ve Fog (1952) okunabilirlik formüllerinde 3+ heceli kelimeler çok heceli olarak belirlenmiştir. Örneğin İngilizcedeki "black" kelimesi beş harfli olup tek heceliyken Türkçedeki "araba" kelimesi beş harfli olup üç heceden oluşmaktadır. Bu durumda da yapılan uygulamalarda metin ve kitapların okunabilirliği düşük çıkmıştır (bkz. Tablo 5-6). Çok heceli kelime sayısı dışında metinlerdeki toplam hece sayısı Ateşman, Fry, Flesch ve Çetinkaya formüllerinde ele alınmıştır. Konuyla ilgili formüllerde sayılar, yabancı kelimeler, sembollerin hece sayısının karakter sayısı ya da okunuşa göre yapılması belirtilmiştir. Bu durumda da ortalama kelime uzunluğu artmıştır (bk. 1. madde). Ayrıca Bezirci ve Yılmaz'a göre (2010) Türkçe metinlerde geçen tek heceli kelimeler metinlerin ortalama kelime uzunluğunu düşürdüğü için metinlerdeki 3, 4, 5, 6 ve daha fazla heceli kelimelerin ortalamaları alınmalıdır.

6. Örneklem Kesiti Sorunu

Problem: Formüllere göre ilk 100, ilk 200, ilk 300'lük kesitler ile baştan, ortadan ve sondan alınan kesitlerin okunabilirlik sonuçları arasında bir farklılık var mı? 100, 200 ya da 300. cümleler cümle ortalarında bitiyorsa ne yapılmalıdır?

Formüllerin bazıları (Dale ve Chall (1948), Fry (1977), Fog (1952), ARI (1967) ve Sönmez (2003) gibi) örneklem olarak ilk 100 ilk 200 ilk 300 gibi bir alanın kullanılması önermiştir. Formüllerin en büyük problemi aynı metin üzerinden denenen okunabilirlik formülünde metin içinde dahi farklı sonuçların ortaya çıkmasıdır. Örneğin; 6. sınıf ders kitabındaki "Meşeler" adlı metnin ilk yarısı ele alındığında

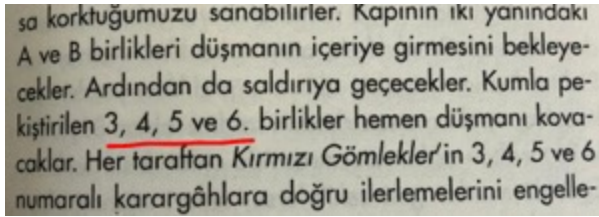
Ateşman'ın formülüne göre 65.20 "Orta Güçlükte" çıkarken ikinci yarısı 77.82 "Kolay" olarak çıkmaktadır. Benzer durum çocuk kitaplarında da ortaya çıkmıştır. *Çıtır Çıtır Felsefe* serisinde ilk 100 kelime ile yapılan denemelerde 84,53 ile çok kolay çıkarken metnin ilk, orta ve son kısmından alınan 100 kelimele kısımlar (toplam 300 kelime) ele alındığında sonuçlar 74,12 kolay olarak çıkmaktadır. Bu durumun oluşmasında ele alınan metinlerin ilk kısımlarının genellikle kısa kelime ve cümle uzunluğuna sahip olması etkilidir. Bu noktada kesitlerin metnin başından, ortasın veya sonunda mı yoksa farklı kısımlar üzerinden alınan parçalar üzerinden mi alınacağı netleştirilmelidir.

Örneklem oluşturulurken alınan ilk 100 gibi kesitler şayet cümle ortasına denk geliyorsa bu durumda yapılan uygulamalarda cümle bitimine kadar alındığında 104, 105 kelimele bir kısım alınmıştır. Cümle alınmadığında ise 97, 96 kelimele bir alan alınmıştır. Her iki sonuç da ölçülmüştür. Bu durumda son cümle alındığında sonuç 6,65 alınmadığında ise 7,19 olmuştur. Bu durum formüllerin seviyelendirilmesi hususunda önem arz etmektedir. Flesch (1948: 228) formül çalışmasında 100. kelimenin cümle ortasına gelmesi durumunda 94, kelimeye kadar geriye gidilebileceği ve 109. kelimeye kadar gidilerek cümlelerin tamamının alınması belirtilmiştir. Fry ise bu formüllerin dışında baştan, ortadan ve sondan 100 kelimele kesitler alınması gerektiğini belirtmiştir. Bu doğrultuda yapılan denemelerde ilk 100 kelimenin alınması gerektiği belirtildiğinde 100. kelimenin cümle ortasına gelmesi durumu okunabilirlik sonuçlarının farklı çıkmasına yol açmıştır.

7. Noktalama işaretleri

Problem: Nokta, virgül, noktalı virgül gibi noktalama işaretlerinin farklı işlevlerinin tanımlanması.

Cümle bitirici olarak Türkçede nokta, ünlem, soru işareti, üç nokta kullanılır. *Pal Sokağı Çocukları* adlı kitaptan alınan aşağıdaki örnekte cümle bitirici olarak kullanılan nokta işaretinin kısaltmalardaki karşılığı yapay zekâ çalışmaları için özel kullanım olarak sisteme tanımlanmalıdır.



Yandaki resimde ise nokta sıra bildirmek amacıyla kullanılmaktadır (64. s.).

8. Parantez içi kullanımlar

Parantez, açıklama yapma, yabancı kelimelere karşılık verme gibi amaçlarla bazen bir kelime bazen de cümle düzeyinde bir veya birden fazla kelime yer alabilmektedir. Yapay zekâyâ parantez tanımlırken nasıl tanımlanmalıdır? Araştırma kapsamında incelenen kitaplarda üç farklı parantez içi kullanılmıştır:

a) Beden dili ve konuşanın hareketlerinin açıklamak için kullanılmıştır:

Örnek: *Çıtır Çıtır Felsefe Serisi, Anlaşmak ve Anlaşmamak*: Lale (telaşla)

b) Yabancı kelimelerin okunuşu için kullanılmıştır:

Örnek: *Pal Sokağı Çocukları*: Einstand (Ayşand)

c) Ara cümleleri belirtmede kullanılmıştır:

Örnek: *Charlie'nin Çikolata Fabrikası*(8. s.): Altı yetişkin (siz yine de bir sayın isterseniz), bir de küçük Charlie...

yor? Bakın, sizinle tanıştığına ne kadar sevindi.
Altı yetişkin (siz yine de bir sayın isterseniz), bir
de küçük Charlie Bucket'tan oluşan bu aile koca ken-
tin kıyısında bir yerde, küçücük bir ahşap kulübede
yaşıyordu.

Konuyla ilgili yapılan denemelerde parantezli ifadeler alındığında 8,19 alınmadığında ise 6,19 sonucu ortaya çıkmıştır. Arada anlamlı bir farkın olmasının yanı sıra yabancı kelimelerin okunuşu da dâhil olmak üzere okuyucunun okuduğu ve bunlara göre zihninde bir tasarım yaptığı düşünüldüğünde belirtilen ifadelerin alınması gerektiği önerilmektedir.

9. Ara söz kullanımları

Ara sözler kitap ya da metinlerde metinle ilgili ayrıntı bilgi vermek için iki kısa çizgi ya da iki virgül arasında kullanılmaktadır. Araştırma kapsamında iki kelimedenden oluşan ara sözlerin yanı sıra çok uzun ara sözlere de rastlanmıştır. Örneğin *Charlie'nin Çikolata Fabrikası* (11-12. s.) adlı kitaptan alınan aşağıdaki ara söz formlerine göre denenmiştir.

şerdi.
Bucketların açlıktan öldükleri söylenemezdi, ama
hepsinin -Charlie'nin dedeleriyle ninelerinin, anne-
siyle babasının, hele küçük Charlie'nin kendisinin-

bütün gün açlık başlarına vurur, mideleri ezim ezim
ezilirdi.

Yapılan denemeler sonrasında ARI formülüne göre ara sözler dâhil edildiğinde sonuç 20,48, dâhil edilmediğinde 14,34 çıkmıştır. Aynı deneme Ateşman'a göre yapıldığında ara sözler dâhil edildiğinde 7,63 Çok Zor, dâhil edilmediğinde 62,53 Orta Güçlükte çıkmıştır. Sonuçlar arasında anlamlı bir farklılık olduğu için bu durumun okunabilirlik formlerinde belirtilmesi ve nasıl ele alınması gerektiği belirtilmiştir.

10. Konuşma balonlarında cümlelerin okunabilirliğe dâhil edilmesi

Karikatürlerin, konuşma balonlarının sayfa kenarlarında sözlerle karşılaşmaktadır. Bu düz metin dışı cümlelerin okunabilirliğe dâhil edilip edilmeyeceği sorun oluşturmaktadır.



Konuyla ilgili kitaplarda yandaki gibi konuşma balonları yer almaktadır (6. s.). Bu kısımlar formlere dâhil edildiğinde 6,68 dâhil edilmediğinde 6,65 sonucu elde edilmiştir. Sonuçlar anlamlı olmasa da konuşma balonlarındaki ifadelerin daha fazla olduğu sayfalara ulaşıldığından okunabilirliğe etki edebileceği düşünülmüştür.

11. Başlıkların okunabilirliğe dâhil edilmesi

Formüllere dâhil edilmesi konusunda tereddüt yaşanan bir sorun da başlıkların alınıp alınmayacağıdır. "Başlıklar alınmalı mı?" sorusuna yönelik yapılan denemelerde başlıklar dâhil edildiğinde ARI formülüne göre sonuç 6,68, başlıklar dâhil edilmediğinde 6,65 sonucu ortaya çıkmıştır. Bazı kitap ve metinlerdeki ara başlıklar sadece tek bir kelimeyken bazılarında *Çıtır Çıtır Felsefe* serisinde olduğu gibi "Herkesin fikri kendine mi?" şeklinde bir cümle şeklidir ancak sonuç anlamlı olarak çıkmadığı için başlıklar sonuçlara dâhil edilmemiştir.

3.3.2. Nitel Kısımla İlgili Problemler

Nitel kısımın alakalı değişkenler sekiz ana başlık altında bulgularda yer almaktadır. Nitel değişkenler bir söz varlığı havuzunun nasıl oluşması gerektiği konusunda bilgi verecektir.

1. Bilinmeyen/zor kelimelerin sayımı

Problem: Bir bilinmeyen kelime, tekrar ettiği kelime sayısı kadar mı ele alınacak?

Bir kelime bilinmiyorsa o kelime kaç kere geçerse geçsin cümlelerin anlaşılabilirliğine etki edecektir. Metin ya da kitaplarda geçen bilinmeyen bir kelime eğer 3 kere geçiyorsa bu bilinmeyen kelimenin 1 kez mi yoksa 3 kez mi alınacağı bir sorun oluşturmaktadır çünkü Dale ve Chall (1948) ile Sönmez (2003) formüllerinde bilinmeyen kelimeler toplam kelime sayısından çıkartılmaktadır. Bu durumda formüllerde bir açıklamaya ulaşamamıştır ancak formüllerde yapılan uygulamalara bakıldığında okunabilirlik sonuçlarının çok düşük çıkmasından dolayı bilinmeyen kelimelerin tekrar ettiği kelime sayısı kadar alınması gerektiği önerilmektedir. Yapay zekânın bu noktada bir kelimenin tekrarını da bir değişken olarak sınaması gerekecek.

2. Deyim ve atasözlerinin bilinmeyen kelime olarak belirlenmesi

Problem: Anlamı bilinmeyen deyim ve atasözleri bir bilinmeyen mi yoksa atasözü ve deyimi oluşturan her bir kelime ayrı ayrı mı bilinmeyen kabul edilecek?

Deyim ve atasözleri çoğunlukla birden fazla kelimedenden oluşur. Bu birliktelikte her bir kelimenin anlamı bilirse de bir araya gelen birlikte ortaya çıkan yeni anlam bilinmeyebilir. Bu noktada bir deyimdeki anlamı bilinen kelime ayrı mı yoksa bir bütün olarak mı bilinmeyen kabul edilecek. Örneğin; *Pembe İncili Kaftan* adlı kitapta geçen "Elçiye zeval olmaz." atasözü tamamı bilinmeyen bir cümledir ve dolayısıyla kaç kelimenin bilinmeyen kelime olarak alınacağı bilinmemektedir. 8. sınıf Türkçe ders kitabında yer alan *Bayrağımızın Altında* adlı metinde yer alan "dikili ağacı kalmamak" deyiminde yer alan kelimeler tek tek değerlendirildiğinde bilinmesine rağmen bir bütün olarak değerlendirildiğin bilinmeyen kelime olarak ya bir ya da üç kelime olarak alınacaktır. Deyimler kelime sayısı olarak alındığında metnin okunabilirliğini düşürmekte tek bir kelime olarak alındığında ise metindeki toplam kelime sayısı ters düşmektedir. Deyimlerin bir özelliği de ikilemelerden oluşabilmesi durumudur. Örneği "kara kara düşünmek" bir deyimdir ancak "kara kara" kısmı da ayrı bir ikileme olarak sayılabilir mi sorusu akıllara gelmektedir. Bu durumda aynı ifadelerin farklı değişkenler içinde alınmasına yönelik bir yönerge geliştirilmelidir. Yapay zekâ bu noktada deyim ve atasözleri havuzu oluşturulması gerekmektedir.

3. Soyut ve mecazlı ifadelerin bilinmeyen kelime olarak ele alınması

Problem: Metinde geçen kelimelerin bir araya gelerek oluşturduğu soyut ya da mecaz anlamlı kelimeler, içerdiği kelime sayısı kadar mı yoksa tek bir kelime olarak mı sayılacak?

Soyut ve mecazlı ifadelerin kaç bilinmeyen kelime olarak alınacağı bir önceki maddeyle aynı problemi oluşturmaktadır. 8. sınıf ders kitabındaki *Bayrağımızın altında* (46. s.) adlı metinde geçen şu cümledeki mecaz ifadedeki kelimeler tek başına bilinen kelimeler olsa da birleştiğinde mecaz anlamda kullanılmaktadır. Bu durumda kaç kelimenin bilinmeyen/zor kelime olarak alınacağı bilinmemektedir:

sahibi kadınla çıktı o da geldi, boynuma sarıldı. Bana öyle geldi ki bu zafer arasında bütün Salihli'deki kadınların kimi sevgili bir vücudu, hepsi yerini yurdunu hatta karnını doyrabilmek güvenini kaybettikleri için arada alınlarından uçan siyah endişe gölgesi Hatice Nine'de yoktu. Oysaki en fakiri, en ihtiyarlı ve en hâlsiziydi. Yanıma gelmek için birkaç defa duvara dayanmıştı. Ama kıymetli, esmer bir kösele

4. Özel adların bilinmeyen kelime olarak alınması

Problem: Anlamı bilinmeyen yabancı özel adları bilinmeyen kelime olarak sayılacak mı?

Özel adların bilinmeyen kelimelere dâhil edilip edilmeyeceği formüllerde belirtilmemiştir. Örneğin yine *Pembe İncili Kaftan* adlı kitapta geçen "Transilvanya, Eflak, Modon, Koron, Santamavro" (183. s.) kelimelerin bilinmeyen kelime olarak alınıp alınmayacağı belirsiz kalmaktadır.

5. Tekrar eden yansıma kelimeler ve ikilemelerin sayımı

Şiirler, okunabilirlik formüllerinin hiçbirinde sınanmamıştır. Bunun sebebi formüllerin kullanıldığı çalışmalarda açıklanmamıştır. Bu kapsamda her bir sınıftan ders kitabındaki ilk şiir metni okunabilirlik formülleriyle denenmiş, şiirler düz yazıya çevrilmemiş ve yaşanan sorular sıralanmıştır.

Tablo 7: Uygulamada Kullanılan Şiirlerin Okunabilirlik Özellikleri

Şiirin adı ve sınıfı	Dale ve Chall	Flesch	Fog	ARI	Smog	Ateşman	Sönmez	Bezirci ve Yılmaz	Çetinkaya
5. sınıf "Memleket İsterim"	Uygulanamıyor çünkü şiir 100 kelimenin altında	32 zor	Uygulanamıyor or çünkü şiir 100 kelimenin altında	8,1 14-15 yaş	Uygulanamıyor or çünkü şiir 30 cümle barındırmıyor.	99,21 Çok Kolay	AO: 0,33 metin anlamsız	3,05 3. sınıf	55 OD: Bağrı OED: 5, 6 sınıf
6. sınıf "Arıyorum"	OD: 1., 2., 3 ve 4. düzey Eğitim aşaması: İlköğretim Yaş: 5-10	17,8 Çok zor	11,3 metin anlaşılabilir	11,98 16-17 yaş	12 eğitim seviyesi	72,65 Kolay	AO: 1,01 anlamsız	10,7 10-11. sınıf	34,11 OD: Eğitimi OED: 8 ve 9 sınıf
7. sınıf "Atatürk'ü Gördüm Düşümde"	Uygulanamıyor çünkü şiir 100 kelimenin altında	30 zor	Uygulanamıyor or çünkü şiir 100 kelimenin altında	7,7 14-15 yaş	Uygulanamıyor or çünkü şiir 30 cümle barındırmıyor.	94,95 Çok Kolay	AO: 0,33 metin anlamsız	3,91 3-4. sınıf	55 OD: Bağrı OED: 5, 6 sınıf
8. sınıf "İnsanla Güzel"	Uygulanamıyor çünkü şiir 100 kelimenin altında	32 zor	Uygulanamıyor or çünkü şiir 100 kelimenin altında	7,5 14-15 yaş	Uygulanamıyor or çünkü şiir 30 cümle barındırmıyor.	93,14 Çok Kolay	AO: 0,33 metin anlamsız	4,10 4. sınıf	55 OD: Bağrı OED: 5, 6 sınıf

Yukarıdaki tabloda da görüldüğü üzere özellikle Türkçe için geliştirilen Ateşman (1997) ve Çetinkaya (2010) okunabilirlik formüllerine göre şiirlerin okunabilirliği yüksek çıkmıştır. Bunun sebebi şiirlerde cümle tanımlamasının bulunmamasından dolayıdır. Bununla birlikte kitabın ya da metnin belirli kısımlarından alınan 100 kelimelik bölümlerin alındığı Smog (1969), Dale ve Chall (1948) ve Fog (1952) testleri üç şiirde uygulanamamıştır. Bunun nedeni ise şiirlerin 30, 58 ve 63 kelimededen oluşmasıdır. Kelime sayısının azlığından dolayı cümle uzunluğunun da kısa çıkması -şiirlerde soyut ifadeler fazla olmasına rağmen- Çetinkaya'nın (2010) formülünde okunabilirliğini seviye olarak düşük çıkarmıştır.

Ayrıca formüllerin genelinde "Kısa cümle, anlaşılır ya da okunabilirliği yüksektir." kabulünden ötürü okunabilirliği yüksek çıkmıştır. Metinlerin okunabilirliği denirken cümle olarak satır değişkeni ele alınmıştır. Bu doğrultuda satırlar cümle olarak değerlendirilmiştir çünkü okunabilirlik formüllerinde cümle olarak alınacak ifadeler ya belirtilmemiş ya da nokta, ünlem, soru işaretiyle biten ifadelerin cümle olarak alınabileceği belirtilmiştir.

SONUÇ

Araştırmanın amacı bilgisayar teknolojilerinin kullanıldığı okunabilirlik formüllerine doğru, geleneksel okunabilirlik formülleri üzerinden eleştiri getirilerek Türkçenin yapısal özelliklerini ortaya koymaktır. Todorascu vd.'ne göre (2013: 11) bir metnin okunabilirliği, ele alınan değişkenlerdeki farklılıklara

dayanır. Taylor'a göre (1947: 19) pedagojik olarak bir araştırma yapılmak isteniyorsa bütün değişkenler bir araya toplanmalıdır. Bu doğrultuda hazırlanacak olan bir sistemde tüm değişkenlerin farklı kullanımlarının betimlemesi gerekmektedir. Dolayısıyla her bir öznel kullanımın tespit edilmesi aynı zamanda Türkçenin ifade gücünün de ortaya konması ve Türkçe okunabilirlik yapay zekâsına veri sağlamada önem arz etmektedir. Bu, yapay zekâ çalışmaları için küçük bir adım olacaktır çünkü modern çalışmalar sadece metinlerdeki farklı değişkenleri ele alarak bir sistem oluşturmaktan öte okuyucunun kim olduğuna, okuma süresine kadar pek çok değişkeni işe koşmaktadır.

Türkçe için en son okunabilirlik formülü 2010 yılında geliştirilmiştir. Dünyada yapay zekâyâ dayalı formül çalışmaları ise 2009 yılı itibarıyla geliştirilmeye başlanmıştır. Okunabilirlik alanındaki bu çalışmalar yapay zekâ öncesi yani geleneksel ve sonrası olarak ikiye ayrılmaktadır. Chen ve Meurers'e göre (2019: 419) geleneksel okunabilir üzerine çalışmalar cümle uzunluğuna odaklı; yeni veya modern çalışmalar ise NLP yani doğal dil işleme ve makine öğrenme (machine learning) üzerine odaklanmaktadır. Geleneksel okunabilirlik çalışmalarından yapay zekâ ile oluşturulmuş modern çalışmalara geçişte François (2015: 79) şimdiye kadar bir metnin diğer metinlere göre daha okunabilir olduğunu belirten özelliklerin hâlâ açık olmadığını, okunabilirliğe yönelik kelime sıklığı, soyut kavram oranı veya cümle türleri gibi değişkenleri ele alan formüllerin geliştirildiğini ancak birleştirilmiş açıklayıcı bir model geliştirilemediğini belirtmiştir. Fry'a göre (2002: 287) okunabilirlik formülleri objektif olarak metnin nasıl puanlanacağı bilgisini verirken içerik olarak belirtilen yaş grubuna uygun mu, metindeki çizimler hikâyeyi yansıtıyor mu, bir sayfada kaç kelime var, okunabilirlik seviyesi öğretim yöntemleriyle ilişkili mi, tekrar eden cümle ve kelimelerin oranı nedir gibi soruların cevabını vermez. Zorbaz (2007: 90) da okunabilirlik formüllerinin metnin yalnızca yapısal okuma güçlük düzeyi hakkında bilgi verdiğini ancak okunabilirliğin metnin nitel özellikleriyle (metnin okuyucuya uygunluğu, anlaşılabilirliği gibi) birlikte alındığında kesin bir sonuca ulaşılabileceğini belirtmiştir.

Formüllere dair;

Okunabilirlik üzerine Türkçede ilk çalışmalar Ateşman (1997) ile başlamıştır. Ateşman (1997: 71) formülünde nicel değişkenleri (cümle ve sözcük uzunluğu) kullanmış "okunabilirlik" ile "anlaşılabilirliğin" iki farklı terim olduğunu belirtmiştir. Çetinkaya (2010: 12) ise nicel değişkenleri okunabilirlik kapsamında olduğunu nitel değişkenlerin ise anlaşılabilirlik kapsamında olduğunu belirtmiş ve okunabilirliğin anlaşılabilirlik için bir ön koşul olduğunu belirtmiştir. Bu durumda aynı kelime sayısına sahip olan cümlelerin, karmaşık dil bilgisi yapısına sahip olsa da benzer okunabilirlik seviyesinde değerlendirilmesine sebep olmuştur. Çoban (2014: 108) da bu duruma dikkat çekmiş, kelime sayısının az olduğu cümlelerin okunabilirlik olarak kolay olarak belirtildiğini bu mantığa göre de "Ben, annem ve babam pazartesi günü Kahramanmaraş'a uzun süredir görüşemediğimiz abimi ziyarete gideceğiz." cümlesinin anlaşılabilirliğinin "Galat-ı meşhur fasih-i mehcurdan evladır." cümlesine göre anlaşılabilirlik olarak daha kolay olması gerekirken tam tersi bir sonuç elde edildiğini belirtmiştir. Araştırmacıya göre metin sadece nicel verilerle ölçülemez, metindeki kelimelerin anlamları da okunabilirlik üzerinde etkilidir. Bu durumu Sönmez (2003: 29) ise çalışmasında Yunus Emre'den verdiği "Beni bende deme, bende değilim. / Bir ben vardır bende, benden içeri. / Süleyman kuş dili bilir dediler. / Süleyman var, Süleyman'dan içeri." dörtlük üzerinden göstermiştir. Sönmez (2003: 29) mevcut okunabilirlik çalışmalarından farklı olarak okunabilirliği "anlaşılabilirlik" olarak ele almış ve çalışmasında bu dörtlüğün Fog testine göre açık ve anlaşılır olmasına rağmen dörtlüğün üniversite öğrencileri tarafından anlaşılmadığını belirtmiştir. Verilen dörtlükte ilgi çeken kısım ise kelimelerin tek başına sık kullanılmasına ve bilinmesine rağmen bir soyut düşüncenin anlatılmasında kullanılması ile okunabilirlik seviyesi düşmüştür. Bu durum da yapılacak çalışmalarda kelimelerin tek başına bir liste hâlinde yer almaması gerektiğini, soyutlamalara dikkat edilmesi gerektiğini göstermiştir.

Türkçede en çok kullanılan okunabilirlik formülü 33 çalışma ile Ateşman'ın (1997) ve 12 çalışmayla Çetinkaya'nın (2010) formülüdür. Türkçe metinler üzerine uygulanan formüllerin çok azının Türkçe üzerine geliştirildiği büyük çoğunluğunun ise yabancı diller için geliştirildiği görülmektedir. Formüllerden bazıları çok işlevsel olmasına rağmen yabancı menşeli olduklarından Türkçenin yapısal özellikler ile örtüşmedikleri için kullanılmadığı anlaşılmıştır. Örneğin çok heceli kelimeleri (3+) içeren okunabilirliği düşük/ileri seviye olarak veren Smog (1969) ve Fog (1952) okunabilirlik testleri de

İngilizce dil yapısına uygun olarak belirlendiği için Türkçe metinlerdeki kelimelerin yarısının çok heceli olarak çıkmasına neden olmuştur. Dale ve Chall'ın (1948) formülü alandaki ilk formüllerden biri olması ve bir derleme dayalı olması onu çok değerli kılsa da kelime havuzunun İngilizce kelimelerden oluşması formülü Türkçe için kullanılamaz kılmaktadır. Fry ise ifade ettiği sonuçları dolayısıyla bu çalışmada kullanılamamıştır. Fry'ın okunabilirlik çizelgesine göre çok heceli kelimelerin toplam kelime sayısına yakın olmasından ve 100 kelime başına düşen cümle az olmasından dolayı sonuçlar Türkçe metinlerde Türkçenin yapısal özelliklerinden dolayı yüksek ya da anlamsız çıkmaktadır.

Araştırmada formüllerin bir metnin okunabilirliğine yönelik aynı sonucu verip vermediği sorgulanmıştır. Aynı metne uygulanan formüllerin okunabilirlik düzeyleri her bir formülde farklı çıkmıştır. Örneğin; 8. sınıf ders kitabındaki *Kaşağı* metni için alınan sonuçlar formüllere göre şu şekildedir: Ateşman: 80.63 kolay, Smog: 13. eğitim seviyesi, Dale ve Chall: 16. düzey okunabilirlik, 22+ yaş, Fog: 15.15 metin zor anlaşılır, Flesch: -19.86 metin çok zor, ARI: 9.60 14-15 yaş, Sönmez: 0.09 metin zor anlaşılır, Çetinkaya: 8. ve 9. sınıf. Smog aynı zamanda çalışmasında Dale-Chall'i de kendi sonuçlarından 2-3 derece fazla çıktığı gerekçesiyle de eleştirmiştir (McLaughlin 1969: 643). Yazıcı ve Temur (2007: 324) da çalışmalarında okunabilirlik formüllerinde yaşanan sorunlardan biri olarak farklı formüllerin aynı kitapta farklı sonuçlar vermesini belirtmişlerdir. Bu sonucun ortaya çıkması, formüllerin çoğunun aynı değişkene sahip olmasına rağmen manidardır. Benzer sonuca ulaşan araştırmacılar bu farklılığı değişik sebeplere bağlamıştır. Anagnostou ve Weir (2006: 9) de çalışmalarında benzer sonuçlara ulaşmış, yazarlar bu durumu formüllerin farklı değişkenlere sahip olmasına bağlamışlardır. Bezirci ve Yılmaz (2010: 54) da yaptıkları çalışmada formüllerin kelimelerdeki hece sayısını ele alması, Türkçede bir kelimenin aynı zamanda bir cümle olabilmesi açısından bu durumun uygun olmaması, tek heceli kelimelerin çok heceli kelimelerle kullanıldığında ortalama kelime uzunluğunu düşürdüğü gerekçeleriyle formüller arasında tutarsızlık olduğunu ve sonuç olarak Türkçe metinlerin okunabilirlik açısından düşük seviyede çıktığını, metinlerin profesyonel okuyucu kitlesine hitap ettiğini belirtmişlerdir. Çakmak ve Çil (2014: 17), Kalın (2017: 985) yaptıkları çalışmalarda hem yabancı hem Türk formülleri ders kitapları üzerinde denemiş ve yabancı formüllerin ders kitapları için ya çok yüksek ya da çok düşük okunabilirlik düzeyi verdiğini belirtmiş, formüllerin Türkçe için uygun olmadığını belirtmiş ancak hangi problemlerin yaşandığını belirtmemişlerdir.

Düzeyleleri bakımından formüller;

Formüllerin çıktıkları veya düzeyleri açısından incelendiğinde bütün formüllerde farklı olduğu tespit edilmiştir. Bazı formüllerde yaş olarak verilirken düzeyler bazı formüllerde sınıf olarak verilmiştir. Ateşman (1997), Fog (1952), Flesch (1948) ve Sönmez (2003) metinleri nicel olarak değerlendirip sonuçlarını soyut bir şekilde ölçmüş seviye ya da yaş olarak hangi okuyucu kitlesine göre kolay/zor, anlaşılabilir/zor anlaşılabilir olduğu belirtilmemiştir. Bu formüller genel amaçlar için hazırlanmış olup farklı alanlarda kullanılmaktadır. Bununla birlikte araştırmacılar tarafından ders kitapları üzerinde uygulanan Ateşman'ın formülünde ortaya çıkan "çok kolay, kolay, orta güçlükte, zor, çok zor" değerlendirmesinin hangi seviye/okuyucu için kolay ya da zor olduğu bilinmemektedir. Bu durumda ele alınan 8. sınıf metninin Ateşman'ın formülüne göre "kolay" çıkması herhangi bir tıp metninin "kolay" çıkmasıyla sadece nicel özellikler değerlendirildiği için ne anlam ifade ettiği bilinmemektedir. Smog (1969), Dale ve Chall (1948), Fry (1977), Çetinkaya (2010) ise eğitim düzeyi (sınıf) ve yaş üzerinden değerlendirme yapmışlardır. Burada çok önemli bilgiyi ARI (Senter ve Smith 1967: 9) formülü üzerinden bulmak mümkündür. Bu formül sadece yaş üzerinden bir ölçüm yapmıştır. Formülü geliştiren Senter ve Smith (1967: 9) çıkan sonuçların okunabilirliğe ilişkin sonuçlar verdiğini ve sonuçların sınıf seviyesine yönelik doğrudan referans veremeyeceğini açıklar. Yazarlara göre bu durumun ortaokul, lise seviyesindeki kitaplardaki metinlerinin seviyenin bir seviye daha üstünde olduğunu ve her öğrencinin aynı dersleri görmediği için farklı seviyelerde olduğunu belirtir. Bu durumda da sınıf seviyesi vermek yerine bir yaş aralığı vermenin daha doğru olduğunu ifade eder. Konuyla ilgili çarpıcı bir açıklamayı ise Taylor (1947) yapar ve pedagojik olarak öğrencilerinin seviyelerinin bir seviye üstünde metinlerin seviye açısından zor olmayacağını belirtir.

Düzyay açısından bazı formüller işe yarar olsa da Türk eğitim sistemine uygun olmadığı için kullanılamamıştır. Örneğin Dale-Chall, Smog ve Fry formüllerini, ABD eğitim sistemine göre metinlerin sınıf seviyelerini belirlemiştir. Dale-Chall'ın formülünde 5-12 yaş arası İlköğretim olarak belirlenmiş, ortaöğretim ise 12 yaştan itibaren uygulanmıştır. Aynı durum Smog okunabilirlik formülünde de yaşanmıştır. Smog formülü 5. sınıftan itibaren başlar ancak 13-16 seviyeler arası lisans, 17-18 arası lisansüstü eğitim seviyesi olarak belirtilmiştir. Bu durum Türk eğitim sistemi için uygun değildir. Fry'ın sonuçları eğitim seviyesi olarak verilmiştir ancak 1. eğitim seviyesinin Türkiye'de hangi sınıfı temsil ettiği bilinmemektedir.

Formüllerdeki nicel değişkenler;

Formüllerin hiçbirinde metnin veya kitabın tamamı incelemeye alınmamaktadır. Formüllerin büyük çoğunluğu ilk 100 kelimeyi inceleme nesnesi olarak kabul etmiştir. Yapılan denemelerde metnin çeşitli kısımları üzerinde uygulanan okunabilirlik sonuçları arasında farklılık olduğu belirlenmiş, ilk 100 kelimelik alanlar üzerinde yapılan denemelerde metinlerin kolay, ortalarından alınan kısımlarında ise daha kolay okunabilirlik sonuçları ortaya çıkmıştır. Benzer şekilde Yazıcı ve Temur (2007: 324) da çalışmalarında bir formülün aynı kitabın farklı kesitlerinde farklı sonuçlar verdiğini belirtmiştir. Redish (2000: 134) de çalışmasında okunabilirlik örneklemini için aynı metin içinde alınan farklı bölümler üzerinde farklı sonuçlar verdiğini tespit etmiştir. "Peki, ne yapılmalı?" sorusuna 1948 yılında Flesch şu cevabı vermiştir: Metnin belli aralıklı kısımlarından örnekler alınmalı. Örneğin; her üç sayfada bir, her üç paragrafta bir paragraf. Smog ve Fry metnin başından, ortasından ve sonundan alınan 100 kelimelik (toplamda 300 kelime) kısımların değerlendirilmesi üzerine gitmiştir. Bu durum ise metinlerin ya da kitapların ilk sayfalarında daha somut ve kısa cümlelerin bulunması ve alınan sonuçların yanıltıcı olabileceği ihtimali kaygısından olabileceği düşünülmektedir. Araştırmada ilk 300 kelimenin değerlendirildiği ya da metnin çeşitli kısımlarından alınan parçaların okunabilirliğinin denemesinin okunabilirlik adına daha anlamlı sonuçlar verdiği tespit edilmiştir.

Okunabilirlik formüllerinin tamamında cümle uzunluğu bir değişken olarak yer almıştır. Formüllere göre cümleler uzunsa metnin okunabilirliği düşüktür ancak bu durumda kısa cümlelerin daha soyut, uzun cümlelerin daha somut olduğu ifadelerle rastlanmıştır. Benzer şekilde Bailin ve Grafstein (2016: 65) de formüllerin bu iddiasına karşı çıkar ve okunabilirlik formüllerinin daha uzun cümlelerin daha karmaşık olduğu görüşüne karşılık cümle uzunluğu ile söz dizimi karmaşıklığının aynı şey olmadığını ve bu ayrımın verilmesi gerektiğini belirtir. Bu araştırmada incelenen Einstein'ın *Özel ve Genel Görelilik Kuramı* (1997) kitabı Çetinkaya'nın (2010) formülüne göre 44 puan olarak eğitsel okuma 8 ve 9. sınıfa uygun olarak çıkmıştır ancak kitapta pek çok soyut ifade (klasik-ussal içeriği devirmek, bir tür kübistik felsefeye dönüştürmek, kübizm de saltık değil ama görelidir, kavramsal nokta boyut kazanarak görgülleştirilir, göreci usdışı ise gerçeğin ve yanlıştın önerisidir vb.) bulunmaktadır.

Cümle değişkeninde; çift tırnak içinde verilen cümlelerin ayrı ayrı alındığında dışta kalan kelimelerin nasıl ele alınacağı, çift tırnak gibi konuşma çizgileriyle verilen cümlelerin nasıl ele alınacağı, cümle bitiricilerin kullanılmadığı cümlelerin nasıl ele alınacağı, sıralı ve bağlı cümlelerin sayımı konusunda sorunlar yaşanmıştır. Yapılan denemelerde aktarma, sıralı, bağlı ve birleşik cümlelerin okunabilirliğinin düşük çıktığı, bu cümlelerin ayrı ayrı sayılmasında ise çeşitli değişkenlerin (çift tırnak dışındaki bir kelime cümle olarak mı ele alınacak, bağlı cümlelerde edat ya da bağlaçların hangi cümleye dâhil edileceği gibi) devreye girdiği anlaşılmıştır. Konuyla ilgili olarak Mirzaoğlu ve Akın'a göre (2015: 5) cümle içinde çift tırnakla belirtilen cümleler de ayrı bir cümle olarak alınmalı, kısa çizgiyle birleştirilen kelimeler dahi ayrı ayrı kelime olarak alınmalıdır. Flesch (1948: 228) çalışmasında edat veya bağlaçlarla birleştirilen cümlelerin parçalanmamasını, aktarım, ünlem veya eksilteli cümlelerin tek bir cümle olarak sayılması gerektiğini ifade etmiştir.

Sayı ve kısaltmalardaki hece sayısının nasıl hesaplanacağına dair farklı uygulamalar tespit edilmiştir. Sayıların kaç kelime olarak alınacağı hem yazıyla hem rakamla yazılan sayıların nasıl değerlendirileceği farklılık göstermektedir. Konuyla alakalı olarak Flesch (1948: 228) sembol (Örneğin, \$ "dolar") ve rakamların okunduğu gibi sayılması gerektiğini ancak bir paragrafta birden fazla çok basamaklı sayılar varsa bunların okunuşunun toplam hece sayısına dâhil edilmemesi gerektiğini belirtmiştir. Fry (1977: 246) ise sayıların birer kelime olarak alınması gerektiğini, her bir rakam ise bir hece olarak kabul

edilmesi gerektiğini ifade eder. Fry'a göre çok basamaklı sayıların okunması zor olduğu için okunabilirlik formüllerinin matematik ders kitaplarında ve sayısal içeriklerin bulunduğu pet şişe bandajlarında vb. uygulanması yanlıştır. Yazara göre okunabilirlik formülleri düz yazılar için üretilmiştir. Ayrıca Çetinkaya (2010) tarih ve sayıların tek kelime olarak kabul edilmesini ancak hece sayısının okunuşa göre sayılması gerektiğini ifade etmiştir. Sayıların yanı sıra kısaltmalarda, parantez içindeki ifadelerde ve ara sözlerin sayımında da benzer durum yaşanmaktadır. Flesch (1948: 228) kısaltmaların ve kısa çizgi ile birleştirilen kelimelerin tek bir kelime olarak sayılması gerektiğini belirtir. Fry (1977: 246) ise kısaltmaların tek bir kelime her bir karakterin ise bir hece olarak kabul edilmesi gerektiğini belirtmiştir. Çetinkaya (2010) ise kısaltmaların hece olarak okunduğu gibi sayılmasını, kelime olarak tek kelime, harf olarak da kısaltma kaç harften oluştuysa o kadar sayılması gerektiğini ifade etmiştir. Araştırmacıların bu açıklamaları ortalama kelime uzunluğunu arttırmakta, cümle uzunluğunu düşürmektedir. Bu durum ise aynı metin içinde yapılan sayı, parantez, kısaltma olan parçaların okunabilirlik seviyesinin düşük, bu değişkenlerin olmadığı parçaların okunabilirlik seviyesinin yüksek olmasına neden olmaktadır. Parantez içindeki ifadelerin ve ara sözlerin okunabilirliğe dâhil olması konusunda ise Lepionka (2003: 112) çalışmasında anahtar içeriğin parantez içlerine alınmamasını ve ara sözlerin mutlaka ele alınması gerektiğini belirtmiştir.

Kelime değişkeninde önemli unsur hece sayısının hesaplanmasıdır. Smog (1969) ve Fog (1952) okunabilirlik formüllerinde 3 ve üzeri heceli kelimeler, çok heceli olarak kabul edilmiştir. Bu formüllere göre çok heceli kelimelere sahip metinler ya okunabilirliği düşük ya da ileri seviye okuyucular için uygun olarak belirlenmiştir. Bu formüllerin dışında Ateşman, Fry, Flesch, Bezirci ve Yılmaz ile Çetinkaya formüllerinde, hece sayısını bir değişken olarak ele almışlardır. Fry (1977: 246) İngilizce metinlerde hece sayısının kelimedeki ünlü harflerin sayısı kadar kabul edileceğini belirtir ancak bu durumda harf sayısının kelime uzunluğundaki belirleyiciliği ortadan kalkmaktadır. Örneğin "black" kelimesi beş harfli olmasına rağmen tek heceli sayılmakta "rhythm" (ritim) kelimesi ise hiç sayılmamaktadır. Türkçede ise en çok 4 harfli bir kelime (Örneğin, Türk) tek heceli olabilmekte "evi" gibi kelimeler iki heceli olabilmektedir. Konuyla ilgili Bezirci ve Yılmaz (2010: 56) Türkçede "bu, şu ve o" zamirleri ile "da, ve, ile, için" bağlaçlarının metinlerde çokça geçtiğini ve bu durumun ortalama kelime uzunluğunu da düşürdüğünü ifade etmişlerdir. Yazarlara göre Ateşman'ın (1998) ve diğer formüllerin okunabilirlik formüllerinde bu istatistiksel farklılık göz ardı edilmiştir. Ancak Bezirci ve Yılmaz'ın (2010) formülü ise kelimelerinin okunabilirliğini hece sayısına göre ele almıştır. Bu durumda hece sayısının artması metinlerin ileri seviyede bir okuyucu kitlesine hitap etmesine neden olmuştur. Bu durumda ise çok heceli kelimeler daha az bilinen kelimeler olmaktadır. Örneğin "apartman" gibi bütün dil kullanıcılarına tarafından bilinen kelimeler çok heceli olmaktadır. Sonuç olarak nicel ve nitel değişkenler arasında uyumsuzluk doğmaktadır. Yazıcı ve Temur (2007: 324) da formüllerin evrensel olmadığını Türkçe ve İngilizcedeki harf, hece, kelime ve cümle uzunlukları farklı olduğu için farklı sonuçlar ortaya çıkardığını belirtmiştir. Türkçe ise eklemeli bir dil özelliğine sahip olduğundan metindeki kelimelerin yarısının çok heceli çıkmasına yol açmıştır. Bu doğrultuda çok heceli kelime kavramı tekrar ele alınmalı, uzun olmasına rağmen harciâlem olan kelimelerin formüllerde nasıl ele alınacağı tekrar gözden geçirilmelidir.

Nitel Değişkenler;

Nitel değişkende en önemli ayrımlardan biri derlemidir. Bir kelime havuzu oluşturularak buradan yapay zekânın kelimelerin bilinmeme oranını tespit etmesi amaçlanmaktadır. "Sistem nasıl işleyecektir?" sorusu burada önemlidir. Yapay zekâyâ sürekli yeni kitaplar yeni metinler tanıtılacak ve sistem belli tekrarı olan kelimeleri gruplayacaktır. Bu oranlar bugün için %80, %60 veya %10'u yoğunluktaki kelimelere göre nicel değişkenlere bir oran atayacaktır. Formüllerden yalnızca Dale ve Chall (1948) ile Sönmez (2003) bilinmeyen/zor kelimelerin sayımını bir değişken olarak formüllerine dâhil etmiştir. Yapay zekâ formülünün teorisyenlerinden François'e göre (2015: 85) Dale-Chall'in formülü yapay zekâ okunabilirlik çalışmalarına benzerlik göstermektedir.

Peki derlem nasıl oluşmalıdır? Bulgularda görüleceği üzere deyimlerin, mecazi kullanımların veya kalıp ifadelerin kullanımında çoğunlukla birden fazla kelime yer almaktadır. Oluşturulacak derlem bir

tane olmayıp birden fazla olmak zorundadır. Sistem bir derlemden bilinmeyen kelimeleri tespit ederken başka bir derlemde deyimleri başka birinden mecazları, atasözlerin almalıdır. Bunun sebebi de en iyi yapay zekâ sistemi de olsa mecazların, atasözleri gibi kullanımlar tek tek tanıtılmalıdır. Bu güncelleme havuzlara ekleme işlemi kullanıcılar tarafından birebir yapılmalıdır. Bundan sonraki aşamada yapay zekâ, taramalarda bu derlemler kullanılabilir. François ve Fairon'e göre (2012: 466) her tür metin kendine özgü özellikler taşır ve hepsi bir arada tek derlem ile değerlendirilemez. Alanına özgü bu formül çalışmasının dışında yapay zekâ adına Fransız SATO-calibrage (<http://www.ling.uqam.ca/ato/>) sistemi bu alanın ilk örneklerinden biridir. Daoust, Laroche ve Ouellet (1996: 206) eğitim bakanlığı ile işbirliği yaparak içinde ders kitaplarının, şiir, şarkı, mektup, davetiye, tekerleme, sözleşme gibi farklı dilsel formatlarının olduğu bir derlem oluşturmuş ve oluşturdukları bu sistemi okuma öğreten öğretmenlerin, dil materyali hazırlayanlarının, yazılım hazırlayan şirketlerin ve üniversitelerin hizmetine sunmuşlardır. Çalışma şu anda da Rizkallah, Daoust ve Plante tarafından devam etmektedir. SATO-calibrage günümüzde sadece Fransızca diline yönelik metin analizi yapabilmekle birlikte eğitim, hukuk, psikiyatri, sosyoloji, medya alanlarındaki metinleri analiz edebilmektedir. Vajjala ve Lucic 2018 yılında "OneStopEnglish corpus" (<https://github.com/nishkalavallabhi/UniversalCEFRScoring>) adında bu sistemi geliştirmişlerdir. "OneStopEnglish corpus" için öncelikle derlem oluştururken yaşanan sorunlar listelenmiş, farklı seviyelerdeki metinler ve bu metinlerdeki bölümler birbirine kıyasla kolay ya da zor olabileceği için bu kısımlar önce karşılaştırılmıştır.

Deyimlerin kaç kelime olarak sayılması gerektiği önemli bir ayrımdır. Nicel formüle göre her kelimenin sayıldığı göz önünde bulundurulduğunda oluşturulacak nitel formülde deyim anlamı dikkate alınarak tek bilinmeyen mi yoksa deyim oluşturulan her bir kelime kadar mı bilinmeyen olacaktır? Deyimlerdeki durum soyutlamalar, atasözleri, yansıma kelimeler, ikilemeler ve duygu ünlemlerinde de yaşanmıştır çünkü bu değişkenlerin hepsinde birden fazla kelime bir araya gelerek tek bir anlamı oluşturmaktadır.

Yabancı kelimelerin yabancı dilde olup özel ad olan kullanımların bilinmeyen kelime olarak sayılması konusunda da bilgi verilmelidir. Yabancı özel adların bilinmeyen kelime olarak alınmasından farklı olarak Fry (1977: 244) çalışmasında özel adların yaş gruplarına göre verilmesi gerektiğini, zor söylenebilen kelimelerin okuyucu tarafından atlandığını belirtmektedir. Araştırmacıya göre *Joe* ile *Joseph* kelimelerinin okuma zorluğu farklıdır. Bu kelimeler de yaş gruplarına göre tercih edilmelidir.

Araştırma kapsamında şiirlerin cümle tanımına uymaması veya noktalama işaretlerinin (nokta, ünlem, soru işareti, üç nokta) şekil açısından cümle bitirici gibi amaçların dışında duyguları ifade etmede bir araç olarak görülmesinden kaynaklı mevcut formüllerle sınınamadığı düşünülmektedir. Mevcut okunabilirlik formüllerinde temel değişken olarak ele alınan cümle değişkeni açıklamasında "bir nokta bitirici ile bitmesi" şeklinde belirtilmiştir. Şiirler yapısı itibarıyla alanyazında "nesre çevrilmeyen anlam..." şeklinde ele alınmış, hatta şair ve yazarlar tarafından nazımların nesre çevrilmesi işinin şiirin tamamen dışında olduğu ve nazım ile nesrin farklı türler olduğu belirtilmiştir (Samsakçı 2012: 2722). Fry (1977: 246) ise çalışmasında şiirlerin okunabilirliğe uygulanmasının yanlış olduğunu belirtmiştir, yazara göre okunabilirlik formülleri düz yazılar için üretilmiştir. Meyer (2003: 270) ise çalışmasında formüllerin 100 kelimedenden az metinlerde sonuç vermediğini, bu doğrultuda şiir, grafik, çizelge, formül, vergi formu ve fatura benzeri metinlere uygulanmaması gerektiğini belirtmiştir.

ÖNERİLER

1. Türkçe yapay zekâ okunabilirlik formülü için oluşturulacak Türkçe derlemi ya da kelime havuzu iki özelliği içermelidir. Birinci olarak kelime sıklığı: yapay zekâ ele aldığı metindeki en çok tekrar eden kelimeleri kolaydan zora sıralamalıdır. Bir kelime çok geçiyorsa bu kelime "bilinen/kolay" kelime olarak etiketlenmeli, az miktarda geçiyorsa "bilinmeyen/zor kelime" olarak etiketlenmelidir. Her kitap tarandığında bu kelime havuzu genişletilmeli, sistem insan müdahalesi olmadan kelimelerin bilinirliğini sıralamalıdır. İkincisi ise kelime havuzunun deyim, atasözü, mecaz ayağıdır. Yapay zekâ bu kısımda iki ya da daha fazla kelimedenden oluşan deyimleri metin içinde görmelidir. Bu durumda kelime havuzuna deyimler oluştukları kelimelerle bir kalıp olarak etiketlenmelidir ancak bu kısımda da kelimelerin özellikle son seslerinde oluşan yumuşama, düşme, değişme gibi ses olayları dikkate alınmalıdır. Bu durumda oluşturulacak derlem/kelime havuzu SATO-calibrage gibi bir değıl,

birden fazla derlemden oluşmalıdır çünkü yapılan okunabilirlik analizi birden fazla derlem aracılığıyla açıklanabilmektedir.

2. Türkçe yapay zekâya dayalı okunabilirlik formülün okuyucu değişkeni üzerine çalışmalar yapılmalı, okuma zamanı gibi değişkenler işe koşulmalıdır.
3. Bu kapsamda bilgisayar uzmanlarıyla disiplinler arası çalışmalar yapılmalıdır.
4. Formülün düzey atamalarında yaş sınıflandırılması önemlidir çünkü ARI formülünde de belirtildiği gibi bir sınıftaki herkes aynı düzeyde olmayabilir. Yaş en azından bir derece daha yakın seviyelendirme için olanak sağlayabilir. İlerleyen aşamalarda kelime havuzunun yaş, cinsiyet, bölge şeklinde çeşitlendirilmesi önem taşımaktadır.
5. Metin türleri de okunabilirlik formülleri altında bir değişken olarak ele alınmalı, nesir dışındaki (şiir, tarifler, kullanım kılavuzları, bloglar, fatura/fiş, hipertext vb.) metin türlerinin de okunabilirliğinin değerlendirilmesine yönelik özellikler de dikkate alınmalıdır.
6. Okunabilirlik üzerine nicel anlamda Türkçede çok fazla çalışma yapılmış olduğu görünse de bu çalışmaların pek çoğunda bir formülün Türkçe metne uygulanması şeklinde gerçekleşmiş olup formüllerin eksiklikleri veya geliştirilmesine dönük herhangi bir önerinin olmaması Türkçe okunabilirlik formülü geliştirme uğraşına bir katkı sağlamamıştır. Bu sebeple ilerleyen yıllarda yapılacak çalışmalarda formüllerin aksayan taraflarının listelenmesi Türkçe okunabilirlik formülünün geliştirilmesinde önemli katkı sağlayacaktır.

KAYNAKÇA

- Anagnostou, N. K. ve Weir, G. R. S. (2006). From Corpus-based Collocation Frequencies to Readability Measure. ICT in The Analysis, Teaching and Learning of Languages, Preprints of The ICTATLL Workshop. Glasgow. 33-46.
- Ateşal, Z. (2016). 8-10. Sınıf "Türkçe ve Türk kültürü" ders kitabının hedef yaş düzeyine uygunluğu. *Turkophone, 1 (1)*, 62-73.
- Ateşman E. (1997). *Türkçede okunabilirliğin ölçülmesi. Dil Dergisi, 58*, 71-74.
- Bağcı, H. ve Ünal, Y. (2013). İlköğretim 8. sınıf Türkçe ders kitaplarındaki metinlerin okunabilirlik düzeyi. *Ana Dili Eğitimi Dergisi, 1 (3)*, 12-28.
- Bailin, A. ve Grafstein, A. (2016). *Readability: text and context*. Macmillan: Hofstra University.
- Baki, Y. (2019). Türkçe dersi 8. sınıf kitabındaki metinlerin okunabilirliği. *Dil Eğitimi ve Araştırmaları Dergisi, 5 (1)*, 30-46.
- Baş, B. ve Yıldız, İ. F. (2015). 2. sınıf Türkçe ders kitabındaki metinlerin okunabilirlik açısından incelenmesi. *Mersin Üniversitesi Eğitim Fakültesi Dergisi, 11 (1)*, 52-61.
- Bezirci, B. ve Yılmaz, A. E. (2010). Metinlerin okunabilirliğinin ölçülmesi üzerine bir yazılım kütüphanesi ve Türkçe için yeni bir okunabilirlik ölçütü. *DEÜ Mühendislik Fakültesi Fen Bilimleri Dergisi, 12 (3)*, 49-62.
- Bozlak, G. Ü. (2018). 2016-2017 ve 2017-2018 eğitim ve öğretim yıllarında 5. sınıf Türkçe ders kitaplarındaki hikâye edici metinlerin Uzun-Çetinkaya formülü ile okunabilirlik düzeyleri/düzeylerinin incelenmesi. *Akra Kültür Sanat ve Edebiyat Dergisi, 6 (14)*, 209-234.
- Büyüköztürk, Ş., Çakmak, K. E., Akgün, E. Ö., Karadeniz, Ş. ve Demirel, F. (2016). *Bilimsel araştırma yöntemleri*. İstanbul: Pegem Akademi Yayıncılık.

- Ceran, D. (2015). Yüz temel eser listesi Türk edebiyatı kategorisinde yer alan kitapların okunabilirlik düzeyleri ile yedinci sınıf öğrencilerinin kitaplar hakkındaki görüşlerinin değerlendirilmesi. *Dil ve Edebiyat Eğitimi Dergisi*, 14, 23-51
- Chen, X. ve Meurers, D. (2019). Linking text readability and learner proficiency using linguistic complexity feature vector distance. *Computer Assisted Language Learning*, 32(4), 418-447.
- Coleman, E. B. (1965). On Understanding Prose: Some Determiners of Its complexity. NSF Final Report GB-2604. Washington, D.C.: National Science Foundation.
- Collins-Thompson, K. (2014). Computational assessment of text readability: A survey of current and future research. *International Journal of Applied Linguistics*, 6, 97-135.
- Çakıroğlu, O. (2015). İlkokul Türkçe ders kitaplarındaki metinlerin okunabilirlik düzeylerinin öğrenme güçlüğü olan öğrenciler açısından değerlendirilmesi. *İlköğretim Online*, 14 (2), 671-681.
- Çakmak, G. ve Çil, E. (2014). 4. sınıf Fen ve teknoloji ders kitabının okunabilirlik formülleriyle değerlendirilmesi: Canlılar dünyasını; gezelim, tanıyalım örneği. *Turkish Journal of Educational Studies*, 1 (3), 1-26.
- Çetinkaya, G. (2010). Türkçe metinlerin okunabilirlik düzeylerinin tanımlanması ve sınıflandırılması (Yayımlanmamış doktora tezi). Ankara Üniversitesi, Ankara.
- Çoban, A. (2014). Okunabilirlik kavramına yönelik bir derleme çalışması. *Dil ve Edebiyat Eğitimi Dergisi*, 9, 96-111.
- Dale, E. ve Chall, J. S. (1948). *A formula for predicting readability*. *Educational Research Bulletin*, 27, 11-20.
- Daoust, L. Laroche ve L. Ouellet. (1996). SATO-CALIBRAGE: Présentation d'un outild'assistance au 'choix et a la rédaction de textes pour l'enseignement. 'Revue québécoise de linguistique', 25 (1), 205-234.
- Dubay, W. H. (2004). *The principles of readability*. ABD: Impact Information.
- Durukan, E. (2014). Metinlerin okunabilirlik düzeyleri ile öğrencilerin okuma becerileri arasındaki ilişki. *Ana Dili Eğitimi Dergisi*, 2 (3), 68-76.
- Einstein, A. (1997). *Özel ve genel görelilik kuramı* (Çev.: Aziz Yardımlı). İstanbul: İdea Yayınevi.
- Elli, T. (2011). İlköğretim 8. sınıf Türkçe ders kitaplarındaki metinlerin anlaşılabilirlik düzeyi (Yayımlanmamış yüksek lisans tezi). Niğde Üniversitesi, Niğde.
- Eraslan, F. (2008). Eğitsel içerikli web sitelerinin okunabilirlik açısından incelenmesi (Yayımlanmamış yüksek lisans tezi). Gazi Üniversitesi, Ankara.
- Erdem, C. (2011). Dil ve anlatım ders kitaplarındaki metinlerin kelime-cümle uzunlukları ve okunabilirlik düzeyleri üzerine bir değerlendirme (Yayımlanmamış yüksek lisans tezi). Gazi Üniversitesi, Ankara.
- Erol, H. F. (2014). Yabancı dil olarak Türkçe öğretiminde temel seviyede kelime edinimi (Yayımlanmamış doktora tezi). İstanbul Üniversitesi, İstanbul.
- Flesch R. (1948). A New Readability Yardstick. *Journal of Applied Psychology*, 32, 221-233.
- François, T. (2015). When readability meets computational linguistics: a new paradigm in readability. *Revue française de linguistique appliquée*, 20 (2), 79-97.
- François, T. ve Fairon, C. (2012). An "AI readability" formula for French as a foreign language. In *Proceedings of the 2012 Conference on Empirical Methods in Natural Language Processing (EMNLP 2012)*, 466-477.
- Fry, E. (1977). Fry's readability graph: Clarification, validity, and extension to level 17. *Journal of Reading*, 21 (3), 242-252.

- Fry, E. (2002). Readability versus leveling. *Reading Teacher*, 56 (3), 286- 291.
- Gunning, R. (1952). *The Technique of Clear Writing*. New York: McGraw-Hill.
- Günhan, E. (2004). Lise düzeyi kimya kitaplarının elektrokimya kısımlarının, fen okuryazarlığı, yanlış kavramlar ve okunabilirlik yönünden analizi (Yayımlanmamış yüksek lisans tezi). Gazi Üniversitesi, Ankara.
- Güven, A. Z., Bal, M. ve Halat, S. (2014). Ortaokul 5. Sınıf Türkçe ders kitaplarında yer alan metinlerin metinsellik ölçütleri açısından uygunluklarının incelenmesi. *Eğitimde Kuram ve Uygulama Dergisi*, 10 (3), 208-231.
- Güven, S. (2010). İlköğretim 6. sınıf fen ve teknoloji ders kitabının okunabilirliği ve hedef yaş düzeyine uygunluğu (Yayımlanmamış yüksek lisans tezi). Selçuk Üniversitesi, Konya.
- Güyer, T., Temur, T. ve Solmaz, E. (2009). Bilgisayar Destekli Okunabilirliği Analizi. *Türk Eğitim Bilimleri Dergisi*, 7(4), 751-766.
- Hızarcı, S. H. (2009). İlköğretim 6. sınıf yeni sosyal bilgiler ders kitaplarının okunabilirlik düzeylerinin incelenmesi (Yayımlanmamış yüksek lisans tezi). Gazi Üniversitesi, Ankara.
- <http://www.indobase.com/study-abroad/countries/usa/usa-education-system.html> (Erişim tarihi: 10.10.2019).
- İskender, E. (2013). Türkçe ders kitaplarındaki metinlerin kelime ve cümle yapılarıyla okunabilirlik düzeyleri arasındaki ilişki. Karadeniz Teknik Üniversitesi, Yüksek Lisans Tezi, Trabzon.
- İşeri, K. (2011). Türkçe ders kitaplarındaki metinlerin metinsellik özellikleri. Editör: Hakan Ülper. *Türkçe Ders Kitabı Çözümlemeler (90-112)*. Ankara: Pegem Akademi Yayıncılık.
- Jones, M. J. (1997). Methodological themes: critical appraisal of the cloze procedure's use in the accounting domain. *Accounting, Auditing and Accountability Journal* 10(1), 105-128.
- Kalın, Ö. U. (2017). Analysis of 7 th grade social studies course book according to different readability formulas. *International Online Journal of Educational Sciences*, 9 (4), 976 -987.
- Kalın, Ö. U. ve Koçoğlu, E. (2017). 6. Sınıf sosyal bilgiler ders kitaplarının farklı okunabilirlik formüllerine göre incelenmesi. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, 17 (4), 2202-2220.
- Karasar, N. (2012). *Bilimsel araştırma yöntemi*. Ankara: Nobel Yayıncılık.
- Karatay, H., Bolat, K. K. ve Güngör, H. (2013). Türkçe ders kitaplarındaki metinlerin okunabilirlik ve anlaşılabilirliği. *The Journal of Academic Social Science Studies*, 6 (6), 603-623.
- Kızılca, H. ve Yılmaz, A. E. (2008). Doğal dilde yazılmış gereksinimlerin analiz yöntemleri ve bu yöntemlerin Türkçe için uygulanabilirliği. *Yazılım Kalitesi ve Yazılım Geliştirme Araçları Sempozyumu Bildiriler Kitabı*, 69-77.
- Klare, G (1975). Assessing Readability. *Reading Research Quarterly*, 10 (1), 62-102.
- Koçağan, F. Ö. (2012). İlköğretim "4-5." sınıf için hazırlanan öyküleyici çocuk kitaplarının okunabilirlik ve değerler yönünden incelenmesi. *Abant İzzet Baysal Üniversitesi, Bolu*.
- Koroğlu, M. ve Zorbaz, K. Z. (2016) Gazi TÖMER yabancılar için Türkçe öğretim setindeki metinlerin kelime-cümle uzunlukları ve okunabilirlik düzeyleri. *TurkishStudies*, 11 (3), 2509-2524.
- Lepionka, M. E. (2003). *Writing and developing your college textbook*. ABD: Atlantic Path Publishing.
- McLaughlin G. H. (1969). SMOG grading-a new readability formula. *Journal of Reading*, 12 (8), 639-646.
- Mert, L. E. (2018). Türkiye'de kullanılan Türkçe ders kitaplarındaki metinlerin okunabilirlikleri. *Mersin Üniversitesi Eğitim Fakültesi Dergisi*, 9 (3), 87-98.

- Meyer, B. J. F. (2003). Text Coherence and Readability. *Topic in Language Disorders*, 23 (3), 204-224.
- Mirzaoğlu, V. ve Akın, E. (2015). 5. sınıf Türkçe ders kitabındaki metinlerin okunabilirliği üzerine bir inceleme. *Sosyal Bilimler Enstitüsü Dergisi*, 5, 146-155.
- Newbold, N. ve Gillam, L. (2010). The Linguistics of Readability: The Next Step for Word Processing. Workshop on Computational Linguistics and Writing: Writing Processes and Authoring Aids (CLandW 2010). June 6, 2010, Los Angeles, 65-72.
- Oakland, T. ve Lane. H. B. (2004). Language, reading, and readability formulas: Implications for developing and adapting tests. *International Journal of Testing*, 4(3), 239-252.
- Öksüz, H. İ. (2019). Okunabilirliğin anlaşılabilirlik üzerindeki etkisinin farklı metin türlerine göre incelenmesi(Yayımlanmamış yüksek lisans tezi). Düzce Üniversitesi, Düzce.
- Özbek, B. A. ve Ergül, C. (2018). İlkokul 4. sınıf ders kitaplarının okunabilirliklerinin değerlendirilmesi. *Mersin Üniversitesi Eğitim Fakültesi Dergisi*, 14 (2), 653-668.
- Özdemir, S. (2016). Beşinci sınıf Türkçe ders kitabındaki öyküleyici ve bilgilendirici metinlerin okunabilirlik durumu. *Sınırsız Eğitim ve Araştırma Dergisi*, 1 (1), 33 - 46.
- Özmen, C. (2019). Yabancı dil olarak Türkçe öğretiminde metin değiştirim teknikleriyle öykülerin yeniden oluşturulması (Yayımlanmamış doktora tezi). Hacettepe Üniversitesi, Ankara.
- Polat, H. (2018). 5. Sınıf Türkçe ders kitabındaki metinlerin okunabilirlik durumu. *Gaziantep University Journal of Educational Sciences*, 2 (1), 1-22.
- Powers, R. D., Sumner, W. A. ve Kearsley, B. E. (1958). Recalculation of four adult readability formulas. *Journal of Educational Psychology*, 49, 99-105
- Qun, L. (2019). Why natural language processing is AI's jewel in the crown. <https://www.huawei.com/en/about-huawei/publications/winwin-magazine/33/why-natural-language-processing-is-ai-jewel-in-the-crown>. Erişim tarihi: 02.03.2020.
- Rada, M. (2016). Kaan Murat Yanık'ın uçurtma mevsimi kitabındaki öykülerin okunabilirlik açısından değerlendirilmesi (Yayımlanmamış yüksek lisans tezi). Sakarya Üniversitesi, Sakarya.
- Raygor, A. L. (1977). The Raygor readability estimate: a quick and easy way to determine difficulty. *Reading: Theory, Research, and Practice*, 259-263.
- Redish, J. (2000). Readability formulas have even more limitations than Klare discusses. *ACM Journal of Computer Documentation*, 24 (3), 132-140.
- Richards, J. C. ve Schmidt, R. (2010). *Longman dictionary of language teaching and applied linguistics*. London: Longman.
- Samsakçı, M. (2012). "Şiir nesir olmayan söz müdür?" Nazım-nesir farkı, şiirde nesirleşme ve mensur şiirin imkânı üzerine. *Turkish Studies*, 7 (4), 2709-2724.
- Senter R. J. ve Smith E. A. (1967). Automated Readability Index. Technical Report, Cincinnati University, Ohio.
- Sönmez, V. (2003). Metinlerin Eğitselliğini Saptamada Matematiksel Bir Yaklaşım (Sönmez Modeli). *Eğitim Araştırmaları*, 10, 24-39.
- Şimşek, E. (2019). Yabancılar Türkçe Öğretiminde Kullanılan Ders Kitaplarındaki Metinlerin Okunabilirlik Düzeyleri Açısından İncelenmesi. Gaziantep Üniversitesi.
- Taylor, W. L. (1957). Cloze readability scores as indices of individual differences in comprehension and aptitude. *Journal of Applied Psychology*, 40 (1),19-26.
- Teke, S. (2016). Çocuk edebiyatı yazarı Hasan Kallıncı'nın öykülerinin okunabilirlik açısından değerlendirilmesi (Yayımlanmamış yüksek lisans tezi). Sakarya Üniversitesi, Sakarya.

- Temizyürek, F. (2010). Türkiye’de okuma seferberliği çerçevesinde seçilen kitapların kelime-cümle uzunlukları ve okunabilirlik açısından incelenmesi. *TÜBAR*, XXVII, 645-654.
- Temur, T. (2002). İlköğretim 5. sınıf Türkçe ders kitaplarında bulunan metinler ile öğrenci kompozisyonlarının okunabilirlik düzeyleri açısından karşılaştırılması (Yayımlanmamış yüksek lisans tezi). Gazi Üniversitesi, Ankara.
- Temur, T., Sarı, M. M. ve Orhon, B. D. (2011). Fen ve sosyal bilimler alanında yapılan okunabilirlik çalışmalarının karşılaştırmalı olarak incelenmesi. *KHO Bilim Dergisi*, 21 (1), 103-121.
- Todirascu, A., François, T., Gala, N., Fairon, C., Ligozat, A., ve Bernhard, D. (2013). Coherence and Cohesion for the Assessment of Text Readability. *Natural Language Processing and Cognitive Science*, 11-19.
- Tosunoğlu, M. ve Özlük, Y. Ö. (2011). Okunabilirlik ve ilköğretim 1. sınıf Türkçe ders kitabındaki düz yazı metinlerinin okunabilirlik açısından değerlendirilmesi. *Millî Eğitim Dergisi*, 189, 219-229.
- Türkben, T. (2019). Readability characteristics of texts in middle school Turkish textbooks. *Educational Policy Analysis and Strategic Research*, 14 (3), 80-105.
- Ülper, H.; Çetinkaya, G. ve Dikici, A. (2018). Examination of factors affecting students’ reading-comprehension achievement with structural equation modeling. *International Journal of Assessment Tools in Education*, 5 (3), 428-442.
- Vajjala, S. ve Lucic, I. (2018). OneStopEnglish corpus: A new corpus for automatic readability assessment and text simplification. In *Proceedings of the Thirteenth Workshop on Innovative Use of NLP for Building Educational Applications*, 297-304.
- Yazıcı, K. ve Temur, T. (2007). Okunabilirlik formüllerinin kullanımına ilişkin yapılan bazı eleştiriler. *EKEV Akademi Dergisi*, 11 (31), 317-324.
- Yılmaz, F. ve Temiz, Ç. (2014). Yabancılarla Türkçe öğretiminde kullanılan ders kitaplarındaki metinlerin okunabilirlik durumları. *International Journal of Language Education and Teaching*, 2, 70-80.
- Ziya, S. (2019). Behiç Ak’ın çocuk romanlarının söz varlığı ve okunabilirlik yönünden değerlendirilmesi (Yayımlanmamış yüksek lisans tezi). Erciyes Üniversitesi, Kayseri.
- Zorbaz, K. Z. (2007). Türkçe ders kitaplarındaki masalların kelime-cümle uzunlukları ve okunabilirlik düzeyleri üzerine bir değerlendirme. *Eğitimde Kuram ve Uygulama*, 3 (1), 87-101.
- Zorlu, K. Ö. (2015). Readability of primary school 3rd grade Turkish and Social Studies text books. *International Journal of Social Sciences and Education Research*, 1 (3), 807-813.

İncelenen Yayınların Kaynakçası

Ders Kitapları

- Baran, Ç. Ş. ve Diren, E. (2018). *Türkçe 5. sınıf ders kitabı*. Ankara: Anıttepe Yayıncılık.
- Demirel, T. (2018). *Türkçe 6. sınıf ders kitabı*. Ankara: Ekoyay Eğitim Yayıncılık.
- MEB. (2018). *Türkçe 6. sınıf ders kitabı*. Ankara: Millî Eğitim Bakanlığı.
- MEB. (2018). *Türkçe 7. sınıf ders kitabı* (Ed. Tolga Kır, Emine Kırmı, Seda Yağız). Ankara: Millî Eğitim Bakanlığı.
- MEB. (2018). *Türkçe 8. sınıf ders kitabı*. Ankara: Millî Eğitim Bakanlığı.

Çocuk Kitapları

Dahl, R. (2014). *Charlie'nin çikolata fabrikası* (Çev.: Celâl Üster). İstanbul: Can Çocuk Yayınları.

İzgi, M. (2012). *Ökkeş dolmuşçu*. İstanbul: Özyürek Yayınları.

Labbé, B. (2018). *Çıtır çıtır felsefe "Anlaşmak ve anlaşılmamak"* (Çev.: Azade Aslan). İstanbul: Günışığı Yayınları.

Molnar, F. (2018). *Pal sokağı çocukları*. İstanbul: Yapı Kredi Yayınları.

Seyfettin, Ö. (2009). *Pembe incili kaftan*. İstanbul: Üç Harf Yayınları.

Akademik Kitap

Einstein, A. (1997). *Özel ve genel görelilik kuramı* (Çev.: Aziz Yardımlı). İstanbul: İdea Yayınevi.

EXTENDED SUMMARY

The research aims to reveal the structural features of Turkish by criticizing the existing formulas towards the readability formulas using computer technologies. Since each subjective use is also an indicator of Turkish expression style, it is important to reveal the expressive power of Turkish and the formation of readability artificial intelligence.

The most used formula among readability formulas is Ateşman's (1997) formula with 33 studies. The formula that followed Ateşman in the second place with 12 studies is the formula of Çetinkaya (2010). When the studies on readability are examined, it is seen that very few formulas are developed in Turkish and foreign languages. Although some formulas are very functional, it has been determined that they are not foreign because they do not match the structural features of Turkish.

After the texts are measured with readability formulas, level information is available to identify the target audience. However, this level of information is not the same in all formulas. It is not known for which level/reader it is easy or difficult to evaluate the "very easy, easy, medium difficulty, difficult, very difficult" evaluation that appeared in the formula of Ateşman applied to textbooks by researchers. Especially, preparing the readability levels of foreign formulas for their education systems makes it difficult to make sense of the results in the studies using these formulas. Dale-Chall, Smog and Fry formulas determined the grade levels of the texts according to the US education system. This makes it difficult to use formulas in educational studies.

In the study, it was questioned whether the formulas give the same result for the readability of a text. The readability levels of the formulas applied to the same text differed from each other. The problems experienced with in the research are discussed under two titles as quantitative and qualitative variables.

At the end of the research, a readability formula should be created by following the Turkish language structure and the characteristics of Turkish. For this, too many books should be read and different uses in these books should be identified one by one and a word pool should be created.