International Journal of Languages' Education and Teaching

E-ISSN: 2148-2705

Volume 14, 2026

Original Research Article

TOMA: Development of Turkish Readability and Text Analysis Software

M. Rahman Kalyoncu[1], Muhammet Memiş[2], Mehmet Kurudayıoğlu[3] & Selçuk Doğan[4]

Abstract

This study aimed to develop a program for the automatic analysis of the readability and cohesion/coherence of Turkish texts and to evaluate the computational accuracy of the program. For this purpose, a program was developed using Python and various NLP tools to automatically calculate readability formulas developed for Turkish as well as a cohesion/coherence formula. The computational accuracy of the program, which calculates more than 15 parameters and eight formulas, was evaluated using a corpus of 57 texts, and the results showed that the program was able to calculate the formulas with a high degree of accuracy. Having demonstrated its high computational accuracy, the program has been made openly accessible to researchers through the link provided in this study. Based on the highly accurate results obtained from the developed program, it is recommended that the TOMA tool be used in large-scale corpus analyses and readability research and that the scope of text analysis/readability studies be expanded by taking advantage of the time savings provided by the program.

Keywords Text Analysis | Readability | Coherence | Cohesion

TOMA: Türkçe Okunabilirlik ve Metin Analizi Programı'nın Geliştirilmesi

Özet

Bu çalışmada Türkçe metinlerin okunabilirliğinin ve bağdaşıklık/tutarlılık durumlarının otomatik olarak analiz edilebilmesi için bir program tasarlanması ve bu programın hesaplama doğruluğunun sınanması amaçlanmıştır. Bu amaç kapsamında Python ve çeşitli NLP araçları kullanılarak Türkçe için geliştirilmiş okunabilirlik formüllerini ve bağdaşıklık/tutarlılık formülünü otomatik olarak hesaplayan program geliştirilmiştir. 15'ten fazla parametreyi ve 8 formülü hesaplayan programın hesaplama doğruluğu 57 metinden oluşan derlem ile değerlendirilmiş ve programın, formülleri yüksek doğrulukla hesaplayabildiği saptanmıştır. Formülleri yüksek doğrulukla hesapladığı teyit edilen program, bu çalışmada verilen bağlantı ile açık erişimli olarak araştırmacıların kullanımına sunulmuştur. Geliştirilen programdan elde edilen yüksek doğruluktaki sonuçlar doğrultusunda TOMA aracının büyük derlem analizlerinde ve okunabilirlik araştırmalarında kullanılmasına ve programın sağladığı zaman tasarrufundan hareketle gerçekleştirilen metin analizi/okunabilirlik çalışmalarının kapsamının genişletilmesine yönelik önerilerde bulunulmuştur.

Anahtar Kelimeler Metin Analizi | Okunabilirlik | Bağdaşıklık | Tutarlılık

Article Info / Makale Bilgisi	*Received / Geliş Tarihi*	*Reviewed / Kabul Tarihi*	*Published / Yayın Tarihi*	*Doi Number / Doi Numarası*
	10.07.2026	29.08.2026	01.09.2026	10.71084/ijlet.1991409

Reference / Kaynakça Kalyoncu, M. R., Memiş, M., Kurudayıoğlu, M., & Doğan, S. (2026). TOMA: Türkçe Okunabilirlik ve Metin Analizi Programı'nın geliştirilmesi. *International Journal of Languages' Education and Teaching*, *14*, 294-309. https://doi.org/10.71084/ijlet.1991409

[1] Arş. Gör., Sinop Üniversitesi, mrkalyoncu@sinop.edu.tr

[2] Doç. Dr., Ondokuz Mayıs Üniversitesi, muhammet.memis@omu.edu.tr

[3] Prof. Dr., Hacettepe Üniversitesi, mkurudayioglu@hacettepe.edu.tr

[4] Öğr. Gör. Dr., Yozgat Bozok Üniversitesi, selcuk.dogan@bozok.edu.tr

Giriş

Okunabilirlik, metinlerin yapısal özelliklerinden kaynaklanan okuma kolaylığı ya da zorluğu olarak ifade edilmektedir. Okuyucuların anlama becerilerinin farklılaşması gibi metinlerin zorlukları da farklılaşmaktadır ve okuma işleminin verimliliği ile anlamanın gerçekleşmesi metnin zorluk/karmaşıklık düzeyiyle okuyucu seviyesinin uyumlu olmasına bağlıdır (Kalyoncu ve Memiş, 2024; Memiş ve Kalyoncu, 2024). Bu açıdan metinlerin okunabilirliği, metinlerin okuyucular tarafından anlaşılması açısından oldukça önemlidir. Hedef kitle ile metnin uyumsuz olması ise günlük hayattan akademik hayata kadar birçok probleme sebep olmaktadır (Kalyoncu, 2025). Bu duruma örnek olarak seviyesinin altında bir metinle karşılaşan öğrencinin motive olamaması, eğitimden yeterli verimi alamaması; seviyesinin üstünde metinlerle karşılaşan öğrencinin ise metni anlayamaması ve motivasyonunun kırılması gösterilebilir (Güneş, 2000). Bu kapsamda hem okuyucu hem de metin özelliklerini içeren bir kavram olarak metin okunabilirliğinin okuma sürecinin en önemli unsurlarından biri olduğu ifade edilebilir.

Metinlerin okunabilirliğinin belirlenmesinde en verimli ve pratik yöntemlerin başında okunabilirlik formülleri gelmektedir. Okunabilirlik kavramının tanımı doğrultusunda okunabilirlik formülleri, metinlerin ölçülebilir yapısal özelliklerini değişken olarak kullanarak metinlerin zorluk düzeyini ve uygun hedef kitleyi saptayan matematiksel denklemlerdir. Dünya üzerinde oldukça uzun süreden beri üretilmiş birçok okunabilirlik formülü olmasına karşın her dilin eşsiz özellikleri bulunduğu için her formül geliştirildiği dile özeldir (Budak, 2005; Çetinkaya, 2010; Dubay, 2004; Klare, 1984).

Türkçe için üretilmiş birçok okunabilirlik formülü bulunmaktadır. Bu formüllerin öncüsü ise 1997 yılında geliştirilen Ateşman formülüdür (Ateşman, 1997). Esasında bu formül, İngilizce için oluşturulmuş Flesch (1943; 1948) formülünün Türkçeye uyarlamasıdır (Kalyoncu ve Memiş, 2024). Öyle ki Flesch (1948) formülünün içeriği 206,835 – 1.015 *(ortalama cümle uzunluğu) – 84.6*(ortalama kelime uzunluğu) şeklindeyken Ateşman formülünün (1997) içeriği 198,825 – 40,175*(ortalama kelime uzunluğu) – 2,610*(ortalama cümle uzunluğu) şeklindedir. Türkçenin eklemeli bir dil olması, hece uzunluğunun metin zorluğunun ölçülmesinde bir değişken olarak kullanılmasını engellese (Kalyoncu, 2025) de Türkçe için geliştirilen sonraki öncü formül olan Çetinkaya formülünde (2010) de hece uzunluğu temel değişken olarak ele alınmıştır. Bu kapsamda ilgili formüllerin kolayca hesaplanabildiği ifade edilebilir. Öte yandan Türkçe için geliştirilmiş bir diğer okunabilirlik formülü olan Bezirci-Yılmaz (2010) formülü ise metindeki sözcükleri hece sayısına göre tasnif etmekte ve cümle başına düşen ortalama "n" heceli sözcük sayısını temel almaktadır. Bu açıdan Bezirci-Yılmaz formülünün de mekanik bir yapıya sahip olduğu ve bireyler tarafından hesaplanmasının diğer formüllere göre görece zor olduğu söylenebilir. Geleneksel okunabilirlik formüllerinin yanı sıra Kalyoncu (2025) tarafından geliştirilen okunabilirlik formülleri ise kompleks cümle oranı, yan cümle oranı, kelimelerin bilinirlik oranı, sözcük çeşitliliği gibi hesaplanması oldukça zahmetli olan birçok değişkeni bünyesinde barındırmaktadır. Kompleks değişkenler metin zorluğunun daha hassas bir şekilde ölçülmesine olanak tanımasının yanı sıra formüllerin hesaplanmasını oldukça güç hâle getirmektedir.

Türkçe için geliştirilmiş okunabilirlik formüllerinin kolay ya da zor hesaplanabilmesinden bağımsız olarak formüllerin büyük derlemler ve hacimli metinler üzerinde uygulanması büyük bir emek gerektirmektedir. Çünkü İngilizce için üretilen okunabilirlik formüllerinden farklı olarak Türkçe için üretilmiş okunabilirlik formüllerinin otomatik olarak hesaplandığı bir platform ya da uygulama bulunmamaktadır. İngilizce için otomatikleştirilmiş metin analiz araçlarına Lexile Analyzer, Coh-Metrix, TAACO gibi araçları örnek göstermek mümkündür. Bu araçlardan Lexile Analyzer, sözcük sıklığı ve cümle uzunluğu değişkenlerini lexile belirtim denkleminde birleştirerek otomatik bir metin güçlüğü puanı üretmektedir. Sözcük sıklığının ölçülmesinde geniş ölçekli bir derlemden yararlanılması da derlem bilgisinin hesaplama sürecine dâhil edilmesini sağlamaktadır (Stenner vd., 2006). Daha kapsamlı bir okunabilirlik aracı konumundaki Coh-Metrix ise sözcük listeleri, sözcük türleri, söz dizimsel ayrıştırıcılar, derlemler ve örtük anlamsal analiz (LSA) gibi nicel dil bilim bileşenlerini esas alarak okunabilirlik, dil ve bağdaşıklık kapsamında birçok ölçütü otomatik olarak hesaplamaktadır (Graesser vd., 2004). Keza bu aracın kullandığı temel bileşenlerden olan LSA derlem temelli bir analiz aracı olarak vektörler arasındaki kosinüs benzerliğini ele almaktadır (Graesser ve McNamara, 2011). Benzer şekilde TAACO ve TAACO 2 araçları da metinlerdeki bağdaşıklık ve tutarlılık düzeyini sözcüksel örtüşme, bağlaç, sözcüksel çeşitlilik, LSA, Word2vec gibi unsurları işe koşarak tespit etmektedir (Crossley vd., 2016; Crossley vd., 2019). Bu modern hesaplama araçlarının yanı sıra İngilizce için geliştirilmiş geleneksel okunabilirlik formülleri (Flesch, 1948; Gunning, 1952; Kincaid vd., 1975; McLaughlin, 1969; Smith ve Senter, 1967) de "readabilityformulas.com" gibi internet siteleri aracılığıyla otomatik olarak hesaplanabilmektedir. Buna karşılık Türkçe için geliştirilen

okunabilirlik formülleri için böyle bir imkândan bahsetmek güç durumdadır. Türkçede okunabilirliğin otomatik hesaplanmasına yönelik tespit edilen ilk içerik "okunabilirlikindeksi.com" başlıklı internet sitesidir. Ancak ilgili internet sitesinde yapılan uygulamalarda doğru sonuçların elde edilemediği görülmektedir. Diğer yandan ulaşılabilen literatürde Türkçe için geliştirilmiş okunabilirlik formüllerinin otomatikleştirildiği bir çalışmaya da rastlanmamıştır. Hâlihazırda Türkçenin eklemeli yapısı da kompleks değişkenlerin otomatikleştirilmesini zorlaştırmaktadır. Öyle ki Türkçede bir sözcük içerisinde bulunduğu cümleye göre farklı yapılarda ortaya çıkmakta, sözcük yapısında değişmezlikten bahsetmek mümkün olmamaktadır (Kalyoncu, 2025). Buradan hareketle sözcük çeşitliliği, kelimelerin bilinirlik oranı gibi değişkenlerin hesaplanabilmesi için öncelikle aynı sözcüklerin farklı çekimlenmiş hâllerinin eşleştirilmesi gerekmektedir.

Açıklanan bağlamda Türkçe için geliştirilmiş okunabilirlik formüllerinin otomatik olarak hesaplanabilmesi büyük önem taşımaktadır. Çünkü daha önce belirtildiği üzere hacimli metinlerin ya da derlemlerin okunabilirlik düzeyinin doğrudan araştırmacılar tarafından kolay ya da hızlı bir şekilde belirlenmesi mümkün değildir. En basit formüllerde dahi metindeki tüm hecelerin ve cümlelerin sayılması gerektiği, ulusal literatürde her yıl onlarca okunabilirlik çalışması gerçekleştirildiği ve bu hesaplamaların araştırmacılar tarafından el yordamıyla yapıldığı dikkate alındığında, bu otomatik hesaplama araçlarına duyulan ihtiyaç daha anlaşılır hâle gelmektedir.

Okunabilirlik formüllerinin otomatik olarak hesaplanmasının yanı sıra metin analizi sürecinde ele alınması gereken unsurlardan biri de metinlerin bağdaşıklık ve tutarlılık düzeyidir. Çünkü bağdaşıklık ve tutarlılık, harfler topluğunun metin olarak kabul edilmesi için taşıması gereken özelliği ifade etmektedir (Aminovna, 2022). Bağdaşıklık, metin yüzeyinde dil bilgisel ve sözcüksel ilişkileri ifade eden ve söylemdeki bazı ögelerin yorumlanmasının bir diğer ögeye bağlı olduğu durumlarda ortaya çıkan bağlantıları (Dascalu vd., 2015; Halliday ve Hasan, 1976) tutarlılık ise metindeki ifadelerin anlamsal ve mantıksal bir bütünlük içermesini (Karatay, 2010; van Dijk, 1981) temsil etmektedir. Bu bağlamda bağdaşıklığın metnin yapısal düzenini, tutarlılığın ise anlamsal düzeni temsil ettiği ve bu unsurların ayrılmaz bir bütün olduğunu söylemek de mümkündür (Leli, 2020).

Metinsellik açısından öneminin yanında bağdaşıklık/tutarlılık kavramları metinlerin okunabilirliğiyle yani metinlerin anlaşılabilme düzeyi ile de yakından ilişkilidir (Graesser vd., 2003; Hall vd., 2014; Ozuru vd., 2009). Çünkü metnin bağdaşıklığı/tutarlılığı, literatürde vurgulandığı üzere metinlerin niteliğini, zorluğunu ve okunabilirliğini etkilemektedir. Türkçe bağlamında ise bağdaşıklığın ve tutarlılığın nesnel olarak hesaplanmasına yönelik yalnızca bir girişim bulunmaktadır. Kalyoncu, Akdoğan, Memiş ve Kamışlı Öztürk (2026) tarafından gerçekleştirilen çalışmada Türkçe için ilk kez bağdaşıklık ve tutarlılığın parametreleri belirlenmeye çalışılmış ve bağdaşıklığın otomatik olarak hesaplanmasına yönelik iki formül oluşturulmuştur. Ancak ilgili formüller de otomatik olarak hesaplanamamakta, değişkenlerin ayrı ayrı tespit edilip hesaplamaların kullanıcılar tarafından gerçekleştirilmesi gerekmektedir.

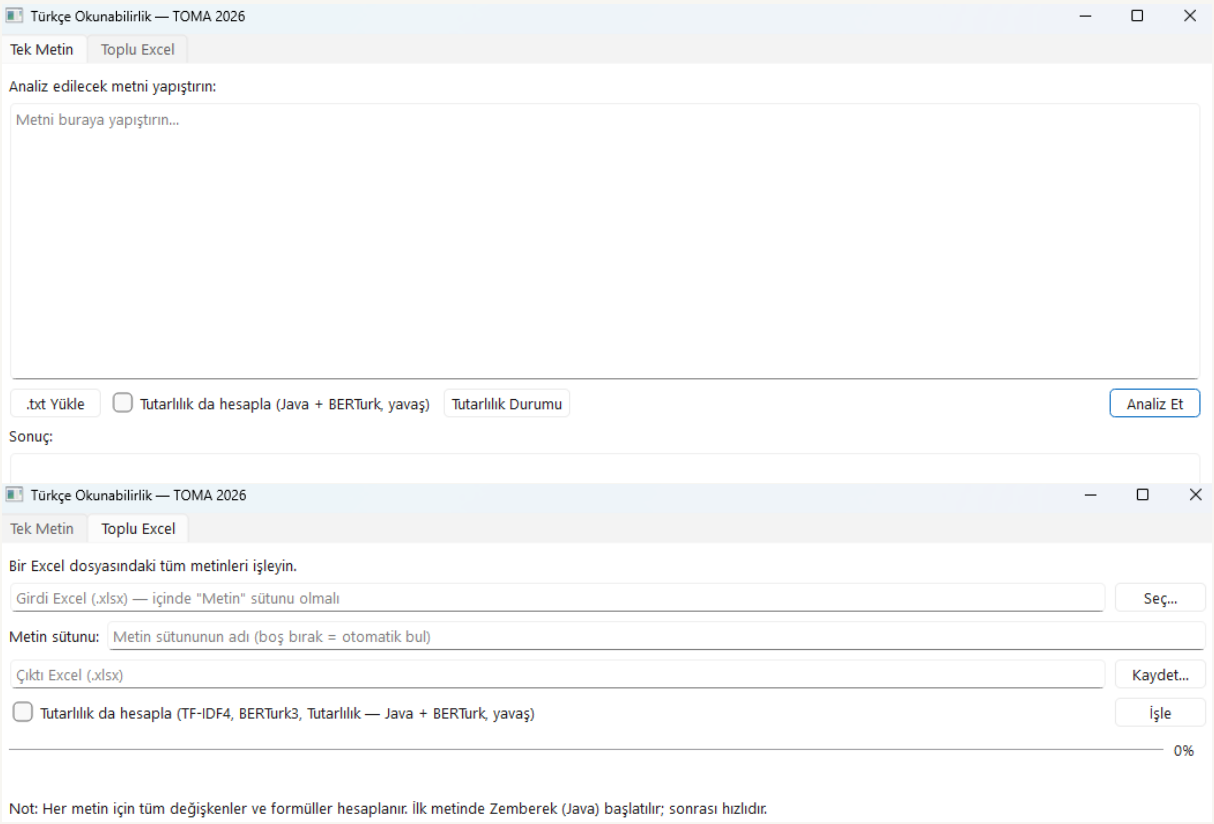
İngilizce gibi diller için metin analizi formüllerinin neredeyse tamamının otomatik olarak hesaplanabildiği dikkate alındığında Türkçe için otomatik analiz araçlarının bulunmaması, okunabilirlik araştırmalarını kısıtlayan bir sınırlılık olarak nitelendirilebilir. Bu bağlamda mevcut çalışma, Türkçe için okunabilirlik formüllerinin ve metin bağdaşıklığı/tutarlılığının otomatik olarak hesaplanmasına yönelik bir araç geliştirmeyi amaçlamaktadır. İlgili aracın geliştirilmesinin yapılan ulusal okunabilirlik çalışmalarında araştırmacılara zaman ve emek tasarrufu sağlaması; ayrıca büyük derlem analizlerini mümkün kılması beklenmektedir.

Yöntem

Araştırma, yazılım geliştirme ve yazılımın doğruluğunu teyit etme doğrultusunda yürütülmüştür. Bu kapsamdaki yazılımın geliştirme süreci esasında araştırmacıların yönlendirmeleri ve kontrolü doğrultusunda "Claude Fable 5" yapay zekâ aracı ile yürütülmüştür. Süreçte, araştırmacılar tarafından AI aracına konu kapsamındaki tüm çalışmalar ve içerikler sunulmuş, değişkenlerin hesaplama metodolojileri aktarılmıştır. TOMA'nın geliştirilmesi sürecinde AI tarafından üretilen sürümlerin çıktıları, araştırmacılar tarafından hesaplanan referans değerlerle karşılaştırılmış; tespit edilen farklılıklar ve farklılıkların kaynakları ele alınmıştır. Bu tespitlerden hareketle özellikle dilsel çözümleme gerektiren değişkenlerde işe koşulan algoritmalar ve NLP araçları araştırmacıların yönlendirmeleri doğrultusunda güçlendirilmiştir. Bu işlem, formüllerin hesaplanmasının mümkün olduğunca doğru bir şekilde uygulanabildiği nihai sürüme erişilinceye kadar sürdürülmüştür. Nihai yazılımın özellikleri ise aşağıda açıklanmıştır:

Genel Mimari

Yazılım mimarisi iki katmandan oluşmaktadır. Birinci katman dil bilimsel hesaplamaları gerçekleştiren bir Python paketinden oluşmaktadır ve bu pakette değişkenler ayrı modüllerde yer almaktadır. Bu işlem ile değişkenlerin bağımsız olarak kontrol edilebilmesi amaçlanmıştır. İkinci katman ise masaüstü arayüzüdür. Arayüz tekil metin işlemesi ve toplu metin işlemesi olmak üzere iki şekilde çalıştırılabilmektedir. Toplu metin işleme süreci ise Excel dosyasına yerleştirilen metinler aracılığıyla gerçekleştirilmektedir. Arayüzde ayrıca bağdaşıklık/tutarlılık hesaplaması isteğe bağlı şekilde sunulmuştur. Bunun nedeni ise ilgili hesaplamaların nispeten uzun bir işlem süresi gerektirmesidir. Arayüze yönelik Görsel 1, aşağıda sunulmuştur:



Görsel 1. Yazılım Arayüzü

Kullanılan Yazılımlar ve Doğal Dil İşleme Araçları

Yazılım temelde Python 3.11 programlama diliyle geliştirilmiştir. Bu süreçte Türkçenin yapısal özellikleriyle uyumlu olan ve açık kaynaklı doğal dil işleme (NLP) araçları işe koşulmuştur. İlgili NLP araçlarını ve yazılımdaki işlevlerini içeren Tablo 1 aşağıda yer almaktadır:

Tablo 1. Programda yazılımlar ve doğal dil işleme araçları

Araç / Yazılım	Sürüm	İşlevi
Python	3.11	Tüm programın temelini oluşturan programlama dili ve çalışma ortamı için tercih edilmiştir.
Zeyrek	0.1.3	KL ve TTR değişkenleri için biçimbilimsel çözümleme sürecinde kullanılmıştır.
Stanza	1.13.0	KCO ve YCO değişkenlerinin hesaplanmasında bağımlılık ayrıştırması ve evrensel bağımlılıklar çözümlemesi için kullanılmıştır.
Zemberek	-	Bağdaşıklık/tutarlılık önişlemesinde cümle sınırlama ve köke indirgeme işlemleri için kullanılmıştır.
jpype1	1.7.1	Python–Java köprüsü; Zemberek'e erişim için kullanılmıştır.
BERTurk	dbmdz/bert-base-turkish-cased	Bağdaşıklık/tutarlılık önişlemesinde bağlamsal cümle gömmeleri için kullanılmıştır.
Transformers	5.12.1	BERTurk modelinin yüklenmesi ve çalıştırılması için kullanılmıştır.

PyTorch	2.12.1	BERTurk modelinin derin öğrenme arka ucu için kullanılmıştır.
scikit-learn	1.9.0	TF-IDF vektörleştirme ve kosinüs benzerliği hesabı için kullanılmıştır.
PySide6	6.11.1	Masaüstü grafik kullanıcı arayüzü için kullanılmıştır.
PyInstaller	6.21.0	Bağımsız çalıştırılabilir Windows uygulaması (.exe) üretimi için kullanılmıştır.
NLTK	3.8.1	Yardımcı dil işleme işlevleri için kullanılmıştır.
pandas / openpyxl	3.0.3	Excel okuma-yazma ve veri işleme için kullanılmıştır
NumPy	2.4.6	Sayısal hesaplamalar için kullanılmıştır
Eclipse Temurin	JDK 17+	Zemberek'in çalışması için gerekli Java çalışma ortamı için kullanılmıştır

Metin Ön İşleme

Yazılımın çalışma sürecinde metin analizinden önce bir dizi önişleme adımı uygulanmaktadır. Bu kapsamda metinler önce cümlelere, ardından sözcüklere ayrıştırılmakta, Türkçeye özgü harfler (örn. â, î, û) standart biçimlere dönüştürülerek (â→a, î→i, û→u) sözcüklerin liste karşılıklarıyla eşleştirilebilmesi sağlanmaktadır. Noktalama ve sayısal belirteçler ise her ölçütün tanımına göre farklı şekilde ele alınmaktadır. Bu işleme örnek olarak kelime listesi temelli hesaplamada salt rakamlardan oluşan belirteçlerin çözümleme dışında bırakılması, yazıyla yazılmış sayı sözcüklerinin ise bilinen sözcükler arasında değerlendirilmesi gösterilebilir. Klasik formüllerde ise hece sayısının doğru hesaplanabilmesi için sayılar, yazılı biçimlere dönüştürülmektedir.

Yazılımda Kullanılan Parametreler ve Formüller

Geliştirilen yazılım, metinlere yönelik 15'ten fazla parametreyi ve 8 farklı formülü analiz etmektedir. Ayrıca analiz edilen formüllere yönelik yorum aralıkları da yazılım çıktısı içinde yer almaktadır. Yazılım bünyesinde hesaplanan parametreler aşağıdaki Tablo 2 içerisinde açıklanmıştır:

Tablo 2. Yazılımda Hesaplanan Parametreler

Sıra	Parametre	İşlev/Özellik
1	Toplam Sözcük Sayısı	Formüllerde kullanılan değişkenlerin oran olarak ifade edilebilmesi için belirlenmektedir.
2	Toplam Cümle Sayısı	OCU, YCO, KCO ve H3-H6 gibi değişkenlerin hesaplanmasında kullanılmaktadır.
3	Özel İsim Sayısı	Metnin içeriğine yönelik bilgi vermesi için eklenmiştir.
4	Bilinir Sözcükler Listesinde Olmayan Sözcük Sayısı	KL değişkeninin hesaplanması için belirlenmektedir.
5	Toplam Hece Sayısı	OSU değişkeninin hesaplanması için belirlenmektedir.
6	Ortalama Cümle Uzunluğu (OCU)	Toplam sözcük sayısının toplam cümle sayısına bölünmesiyle bulunmaktadır.
7	Ortalama Sözcük Uzunluğu (OSU)	Toplam hece sayısının toplam sözcük sayısına bölünmesiyle hesaplanmaktadır.
8	Kelime Listesi Puanı (KL)	Listede olmayan sözcük sayısının toplam sözcük sayısına bölünmesi ve 100 ile çarpılmasıyla hesaplanmaktadır.
9	Type/Token Ratio (TTR)	Metindeki benzersiz gövde sözcük sayısının toplam sözcük sayısına bölünmesi ve 100 ile çarpılmasıyla belirlenmektedir.
10	MATTR	500 sözcükten uzun metinlerde TTR verisinin 500 sözcüklük örneklemle metinde devamlı olarak hesaplanması ve ortalama alınmasıyla hesaplanmaktadır.
11	KCO	Basit olmayan cümle sayısının toplam cümle sayısına bölünmesi ve 100 ile çarpılmasıyla belirlenmektedir.
12	YCO	Toplam fiilimsi sayısının toplam cümle sayısına bölünmesi ve 10 ile çarpılmasıyla bulunmaktadır.
13	TF-IDF	Bağdaşıklık/tutarlılık değerinin hesaplanmasında kullanılmaktadır.
14	BERTurk	Bağdaşıklık/tutarlılık değerinin hesaplanmasında kullanılmaktadır.
15	H3, H4, H5 ve H6	Bezirci-Yılmaz formülünün hesaplanmasında kullanılmaktadır.

Yazılımda hesaplanan parametreler esasında formül değerlerinin belirlenmesinde kullanılmaktadır. Yazılımın hesaplayabildiği 8 formülün içeriği ise aşağıdaki Tablo 3 içerisinde yer almaktadır:

Tablo 3. Yazılımın Hesapladığı Formüller

Formül İsmi	Denklem
Ateşman	198,825 – 40,175(OSU) – 2,610(OCU)
Çetinkaya	118,823 – (25,987 x OSU) – (0,971 x OCU)
Bezirci-Yılmaz	$\sqrt{OCU\ X\ [(H3\ X\ 0{,}84) + (H4\ X\ 1{,}5) + (H5\ X\ 3{,}5) + (H6\ X\ 26{,}35)]}$
K1	(-29,829) + (,508*KL) + (1,191*TTR) + (,733*KCO) – (,66*OCU)
K2	25,457 + (,877*KL) + (,668*TTR) + (,333*OCU)
K3	25,153 + (,834*KL) + (,686*TTR) + (,226*YCO)
K4	43,368 + (1,008*KL) + (,624*OCU)
Bağdaşıklık/Tutarlılık	(12,318) + (25,386[TF-IDF]) + (–9,651[BERTurk])

Tablo 3 kapsamında bağdaşıklık/tutarlılık formülünün ele alınmasında fayda vardır. Çünkü gerçekleştirilen araştırmada (Kalyoncu vd., 2026) esasında iki farklı bağdaşıklık/tutarlılık formülü üretilmiştir. Bu formüllerden birincisi NLP unsurlarına ek olarak metindeki bağlaç ve zamir oranlarını da içermektedir. Ancak zamirlerin hesaplanması sürecinde, işaret zamirlerinin işaret sıfatlarından mekanik bir şekilde kolayca ayırt edilememesinden dolayı kapsamı daha dar olan ikinci formül kullanılmıştır.

Hesaplama Doğruluğunun Saptanması

Yazılımın hesaplama doğruluğunun sorgulanması için 17.678 sözcükten ve 57 metinden oluşan bir derlem kullanılmıştır. Derlemin belirlenmesi sürecinde derlemin heterojen bir yapıda olması, farklı amaçlarla ve farklı dönemlerde oluşturulmuş metinleri yansıtmasına özen gösterilmiştir. Böylece aykırı örnekleme yöntemi tercih edilmiş ve yazılımın farklı yapılardaki metinlerde de geçerli sonuçlar vermesinin sağlanması amaçlanmıştır. Bu kapsamda kullanılan metinlerin özellikleri Tablo 4'te sunulmuştur:

Tablo 4. Derlem Özellikleri

Özellik	n / M	SS	Min.	Maks.
Metin sayısı	57	–	–	–
Sözcük sayısı	310,14	122,85	89	686
Ortalama cümle uzunluğu	10,18	5,93	4,22	37,60
Ortalama sözcük uzunluğu	2,67	0,20	2,33	3,29
Hece sayısı	829,18	330,52	242	1750
Kelime listesi puanı (KL)	8,54	9,82	0,00	42,50
Type/Token Ratio (TTR)	36,26	9,10	22,14	72,04
Karmaşık cümle oranı (KCO)	57,25	26,54	9,68	100,00
Yan cümlecik oranı (YCO)	11,68	12,19	0,00	61,00
TF-IDF	0,077	0,030	0,009	0,151
BERTurk	0,800	0,119	0,390	0,971
Bağdaşıklık/Tutarlılık	6,54	1,44	3,94	9,56

Tablo 4'te nicel değerleri sunulan metinlerin ortalama uzunluğu 310,14 sözcüktür (SS = 122,85) ve metin uzunlukları 89 ila 686 sözcük arasında değişmektedir. Derlemdeki metinlerin 11'i (%19,3) 200 sözcükten kısa, 13'ü (%22,8) 200-299 sözcük, 23'ü (%40,4) 300-399 sözcük ve 10'u (%17,5) 400 ve üzeri sözcük uzunluğundadır. Bunun yanında derlemdeki metinlerin 19'u bilgilendirici/açıklayıcı, 15'i öyküleyici/edebî, 11'i diyalog/söyleşi, 4'ü otobiyografik, 3'ü akademik/teknik, 3'ü tarihî/arkaik, 2'si sözlü anlatım niteliği taşımaktadır. Bu kapsamda derlemdeki metinlerin oldukça heterojen bir yapıda olduğu ve Türkçenin farklı dil kullanım özelliklerini yansıttığı ifade edilebilir.

Derlemdeki metinlerin okunabilirlik ve bağdaşıklık/tutarlılık değerleri henüz yazılım geliştirilmeden önce belirlenmiştir. Bu süreç, okunabilirlik literatürüne yönelik akademik çalışmaları olan ve formül geliştirme sürecini deneyimlemiş bir uzman tarafından gerçekleştirilmiştir. Formüller, öznel yorumlardan uzak bir şekilde yapılandırıldığı ve matematiksel işlemlere dayandığı için bu süreçte birden fazla uzman puanlaması gerçekleştirilmemiştir. Ardından metinler geliştirilen yazılımla analiz edilmiş ve yazılımdan elde edilen sonuçlar ile araştırmacılar tarafından belirlenen okunabilirlik ve bağdaşıklık/tutarlılık değerleri arasındaki korelasyonlar

Pearson korelasyon analizi ile incelenmiştir. Bu süreçte öncelikle saptamaların normallik değerleri ele alınmış ve değerlerin ± 2 aralığında yer aldığı ve normal dağılıma sahip olduğu (George ve Mallery, 2010) görülmüştür.

Analiz sürecinde korelasyon analizleri tek başına hesaplama doğruluğunu garanti etmediği için ek olarak normalize edilmiş ortalama mutlak hata değerleri, intraclass korelasyon analizi, Bland-Altman analizi de yazılım sonuçlarının doğruluğunun belirlenmesinde işe koşulmuştur. ICC analizi iki yönlü karma etkiler modeli (Two-Way Mixed Effects), mutlak uyum tanımı (Absolute Agreement) ve tek ölçüm (Single Measures) esas alınarak gerçekleştirilmiş ve %95 güven aralıklarıyla birlikte raporlanmıştır. Bland-Altman analizinin gerçekleştirilmesi sürecinde ise farklar TOMA sonuçlarından referans değerlerin çıkarılmasıyla hesaplanmış; sistematik sapma ortalama fark (bias), %95 uyum sınırları ise ortalama fark ± 1,96 standart sapma olarak belirlenmiştir. Süreçte ayrıca normalize edilmiş ortalama mutlak hata değerleri "(|Referans Değer – Ölçülen Değer|) / (Ölçüm aralığındaki maksimum değer - Ölçüm aralığındaki minimum değer) * 100" hesaplamasıyla belirlenmiştir. Analizler SPSS 25 programıyla gerçekleştirilmiştir. Analizlerin gerçekleştirilmesiyle birlikte aracın ürettiği değerlerin ölçüt değerlerle yüksek korelasyon ve düşük hata göstermesi, yazılımın formülleri yüksek doğrulukla otomatik olarak uyguladığının bir göstergesi olarak kabul edilmiştir.

Bulgular

Araştırma sürecinin bulguları, geliştirilen yazılımın hesaplama doğruluğunun irdelenmesini içermektedir. Bu kapsamda ilk olarak yazılımda bulunan parametrelerin ölçüm doğruluğu ele alınmıştır. Bu doğrultuda 57 metinden oluşan derlemde araştırmacılar tarafından tespit edilen parametre değerleri ile yazılımın hesapladığı değerler arasındaki korelasyon incelenmiştir. Sonuçlar Tablo 5 içerisinde yer almaktadır:

Tablo 5. Formüllerin Hesaplanmasında Kullanılan Değişkenlerin Korelasyon Değerleri

Değişken	N	Korelasyon	%95 GA	p		M	SS
Sözcük Sayısı	57	,998**	[,997-,999]	<,001	Yazılım	311	124
					Orijinal	310	122
Cümle Sayısı	57	,987**	[,978-,993]	<,001	Yazılım	35	18
					Orijinal	36	20
Hece Sayısı	57	1,000**	[,999-,999]	<,001	Yazılım	830	331
					Orijinal	829	330
OSU	57	,975**	[,957-,985]	<,001	Yazılım	2,6	,208
					Orijinal	2,6	,204
KL	57	,987**	[,977-,992]	<,001	Yazılım	8,9	10
					Orijinal	8,5	9,8
TTR	57	,958**	[,930-,975]	<,001	Yazılım	37	9,1
					Orijinal	36	9,1
OCU	57	,992**	[,987-,996]	<,001	Yazılım	10	5,9
					Orijinal	10	5,9
KCO	57	,962**	[,936-,977]	<,001	Yazılım	59	26
					Orijinal	57	26
YCO	57	,952**	[,919-,971]	<,001	Yazılım	10	10
					Orijinal	11	12
TF-IDF	57	1,000**	[,999-,999]	<,001	Yazılım	,07	,03
					Orijinal	,07	,03
BERTurk	57	1,000**	[,999-,999]	<,001	Yazılım	,80	,11
					Orijinal	,80	,11

Tablo 5 incelendiğinde yazılım tarafından belirlenen ve formüllerin hesaplanmasında kullanılan tüm parametrelerin uzman değerlendirmeleriyle oldukça yüksek düzeyde tutarlılık gösterdiği dikkat çekmektedir. Öyle ki toplam hece sayısı (r = 1,000), TF-IDF (r = 1,000) ve BERTurk (r = 1,000) parametrelerinde korelasyon katsayısının mükemmel düzeyde, toplam sözcük sayısı (r = ,998), ortalama cümle uzunluğu (r = ,992), kelime listesi puanı (r = ,987) ve toplam cümle sayısı (r = ,987) değişkenlerinin katsayısının ise çok yüksek düzeyde olduğu saptanmıştır. Görece düşük korelasyon katsayıları ise YCO (r = ,952) ve TTR (r = ,958) parametrelerinde gözlenmiş, ancak bu değerler de yüksek düzeyde ilişki sınırları içerisinde kalmıştır. Parametrelerin ortalamaları ve standart

sapma değerleri ele alındığında ise yazılım ve araştırmacı hesaplamalarının birbirine çok yakın merkezi eğilime ve dağılım değerlerine sahip olduğu görülmektedir.

Parametrelerin tutarlılığının incelenmesinin ardından formüllerin tutarlılığı ele alınmıştır. Bu sürece yönelik bulgular Tablo 6'da sunulmuştur:

Tablo 6. Formüllere Yönelik Korelasyon Değerleri

Değişken	N	Korelasyon	%95 GA	p		M	SS
Ateşman	57	,997**	[,995-,998]	<,001	Yazılım	64	20
					Orijinal	64	20
Çetinkaya	57	,998**	[,997-,999]	<,001	Yazılım	39	9,2
					Orijinal	39	9,4
Bezirci-Yılmaz	18	,999**	[,997-,999]	<,001	Yazılım	16	10,7
					Orijinal	16	11,1
Bağdaşıklık/Tutarlılık	57	1,000**	[,999-,999]	<,001	Yazılım	6,5	1,4
					Orijinal	6,5	1,4
K1	57	,964**	[,940-,979]	<,001	Yazılım	56	21
					Orijinal	52	22
K2	57	,981**	[,967-,989]	<,001	Yazılım	62	12
					Orijinal	60	12
K3	57	,977**	[,961-,986]	<,001	Yazılım	61	11
					Orijinal	59	12
K4	57	,991**	[,984-,995]	<,001	Yazılım	58	12
					Orijinal	58	12

Tablo 6 ele alındığında tıpkı Tablo 5'te olduğu gibi yüksek korelasyon değerleri dikkat çekmektedir. Öyle ki Ateşman (r = ,997), Çetinkaya (r = ,998), Bezirci-Yılmaz (r = ,999) ve Bağdaşıklık/Tutarlılık (r = 1,000) formüllerinin tamamında yazılım ile araştırmacının hesaplamaları arasında çok yüksek düzeyde anlamlı ilişki bulunmuştur. Kalyoncu (2025) tarafından üretilen formüllerin (K1, K2, K3, K4) korelasyon değerleri ise ,964 ile ,991 arasında değişmekte olup benzer şekilde yüksek düzeydedir.

Gerçekleştirilen korelasyon analizleri, yazılım hesaplamaları ile araştırmacıların hesaplamaları arasındaki doğrusallığı ortaya koymakla birlikte iki hesaplama yöntemiyle elde edilen sonuçların birbirine ne kadar yakın olduğunu saptayamamaktadır. Bu bağlamda farklı ölçümlerin yakınlığını ele almak için intraclass korelasyon analizi, normalize edilmiş mutlak hata ve hesaplama sapması büyüklüğü incelenmiştir. Gerçekleştirilen analize ilişkin sonuçlar aşağıdaki Tablo 7'de sunulmuştur.

Tablo 7. Mutlak Uyum ve Hata Oranları

V	N	NMAE (%)	ICC	%95 GA (L)	%95 GA (U)
Toplam Sözcük	57	0,31	0,998	,997	,999
Toplam Cümle	57	2,27	0,982	,963	,990
Toplam Hece	57	0,06	1,000	1,000	1,000
OSU	57	1,48	0,974	,956	,985
KL	57	2,61	0,986	,976	,992
TTR	57	4,14	0,945	,860	,974
OCU	57	1,36	0,991	,983	,995
KCO	57	6,19	0,959	,929	,976
YCO	57	3,22	0,936	,892	,962
TF-IDF	57	0,01	1,000	1,000	1,000
BERTurk	57	0,00	1,000	1,000	1,000
K1	57	5,77	0,951	,874	,977
K2	57	3,48	0,973	,922	,988
K3	57	3,64	0,971	,941	,985
K4	57	2,33	0,990	,981	,994
Ateşman Puanı	57	0,93	0,996	,993	,998
Çetinkaya Puanı	57	0,85	0,998	,996	,999
Bezirci-Yılmaz Puanı	18	1,30	0,998	,995	,999
Bağdaşıklık/Tutarlılık	57	0,04	1,000	1,000	1,000

Tablo 7 kapsamında ilk olarak normalize edilmiş mutlak hata (NMAE [%]) değerleri ele alındığında toplam sözcük, hece, TF-IDF, BERTurk, Ateşman, Çetinkaya ve Bağdaşıklık/Tutarlılık ölçümlerinin mutlak hata yüzdesinin %1'den az olduğu görülmektedir. Bunun yanı sıra toplam cümle, OSU, KL, TTR, YCO, K2, K3, K4, Bezirci-Yılmaz ölçümlerinin ise mutlak hata yüzdesinin %5'ten az olduğu dikkat çekmektedir. Diğer yandan yazılımın ölçümlerinde KCO (%6,19) ve K1(%5,77) ölçümlerinin hata oranının görece yüksek olduğu görülmektedir. Bu kapsamda KCO ve K1 ölçümlerinin temkinli bir şekilde yorumlanması, gerekli durumlarda araştırmacılar tarafından doğrulama gerçekleştirilmesi gerektiği, diğer ölçümlerin ise yüksek düzeyde doğru sonuçlar verdiği ifade edilebilir.

Normalize edilmiş mutlak hata değerlerinin yanında ICC değerlerinin ,936 ile 1,000 arasında değiştiğini ve tüm ölçümlerde mükemmele yakın düzeyde mutlak uyum bulunduğunu söylemek mümkündür. Öyle ki TF-IDF, BERTurk, Tutarlılık ve Toplam Hece değişkenlerinde ICC değeri 1,000'e ulaşmış, NMAE oranı ise %0,1'in altında kalmıştır. Bu durum söz konusu değişkenlerde yazılımın araştırmacının hesaplamalarıyla pratik olarak birebir örtüştüğünü göstermektedir. Buna karşın KCO (ICC = ,959) ve K1 (ICC = ,951) değişkenlerinde diğer değişkenlere kıyasla görece daha düşük uyum görülse de ICC değerinin ,95'in üzerinde kalması, mutlak uyumun yüksek düzeyde korunduğuna işaret etmektedir. Nitekim 19 değişkenin 17'sinde NMAE oranının %5'in altında kalmasından ve ICC değerlerinin tüm değişkenlerde ,93'ün üzerinde olmasından hareketle geliştirilen yazılımın ilgili parametre ve formülleri yüksek düzeyde doğruluk ve mutlak uyumla otomatik olarak hesaplayabildiği söylenebilir. Bahsi geçen doğruluk değerlerinin incelenmesinin ardından Bland-Altman analizi gerçekleştirilmiştir. İlgili analizin sonuçları Tablo 8'de sunulmuştur:

Tablo 8. Bland-Altman Analizi Sonuçları

Değişken	**N**	**Bias**	**%95 Uyum Sınırları**
Toplam Sözcük	57	1,67	[-12,33; 15,67]
Toplam Cümle	57	-1,49	[-8,19; 5,21]
Toplam Hece	57	0,96	[-6,46; 8,39]
OSU	57	-0,009	[-0,100; 0,083]
KL	57	0,42	[-2,79; 3,63]
TTR	57	1,56	[-3,62; 6,73]
OCU	57	0,30	[-1,15; 1,74]
KCO	57	2,20	[-12,10; 16,50]
YCO	57	-1,07	[-8,80; 6,66]
TF-IDF	57	-0,000002	[-0,00; 0,00]
BERTurk	57	0,000004	[-0,00; 0,00]
K1	57	3,56	[-8,28; 15,41]
K2	57	1,55	[-3,24; 6,34]
K3	57	1,22	[-4,08; 6,52]
K4	57	0,61	[-2,75; 3,97]
Ateşman	57	-0,58	[-3,80; 2,63]
Çetinkaya	57	-0,22	[-1,42; 0,98]
Bezirci-Yılmaz	18	0,20	[-1,07; 1,48]
Bağdaşıklık/Tutarlılık	57	0,000249	[-0,00; 0,00]

Tablo 8'deki bulgular, genel olarak Bland-Altman analizi sonuçlarının normalize edilmiş mutlak hata sonuçlarıyla uyumlu olduğunu ve TOMA tarafından üretilen değerler ile referans hesaplamalar arasındaki ortalama farkların genel olarak düşük olduğunu göstermektedir. Öyle ki bağdaşıklık/tutarlılık, OSU, KL, OCU, Ateşman ve Çetinkaya ölçümlerinde sistematik sapma düşük durumdadır. Buna karşılık KCO ve K1 ölçümlerinin diğer değişkenlere göre daha geniş uyum sınırlarına sahip olduğu dikkat çekmektedir. Bunun yanında TOMA'nın K1, K2, K3 ve K4 formüllerini bir ölçüde yüksek, Ateşman ve Çetinkaya formüllerini ise bir ölçüde düşük hesapladığı görülmektedir. Bu bağlamda TOMA hesaplamaları ile referans hesaplamalar arasındaki sistematik farkın genellikle düşük olduğunu, ancak KCO ve K1 gibi karmaşık hesaplamalarda sapmaların daha yüksek olabileceğini ifade etmek mümkündür.

Sonuç ve Öneriler

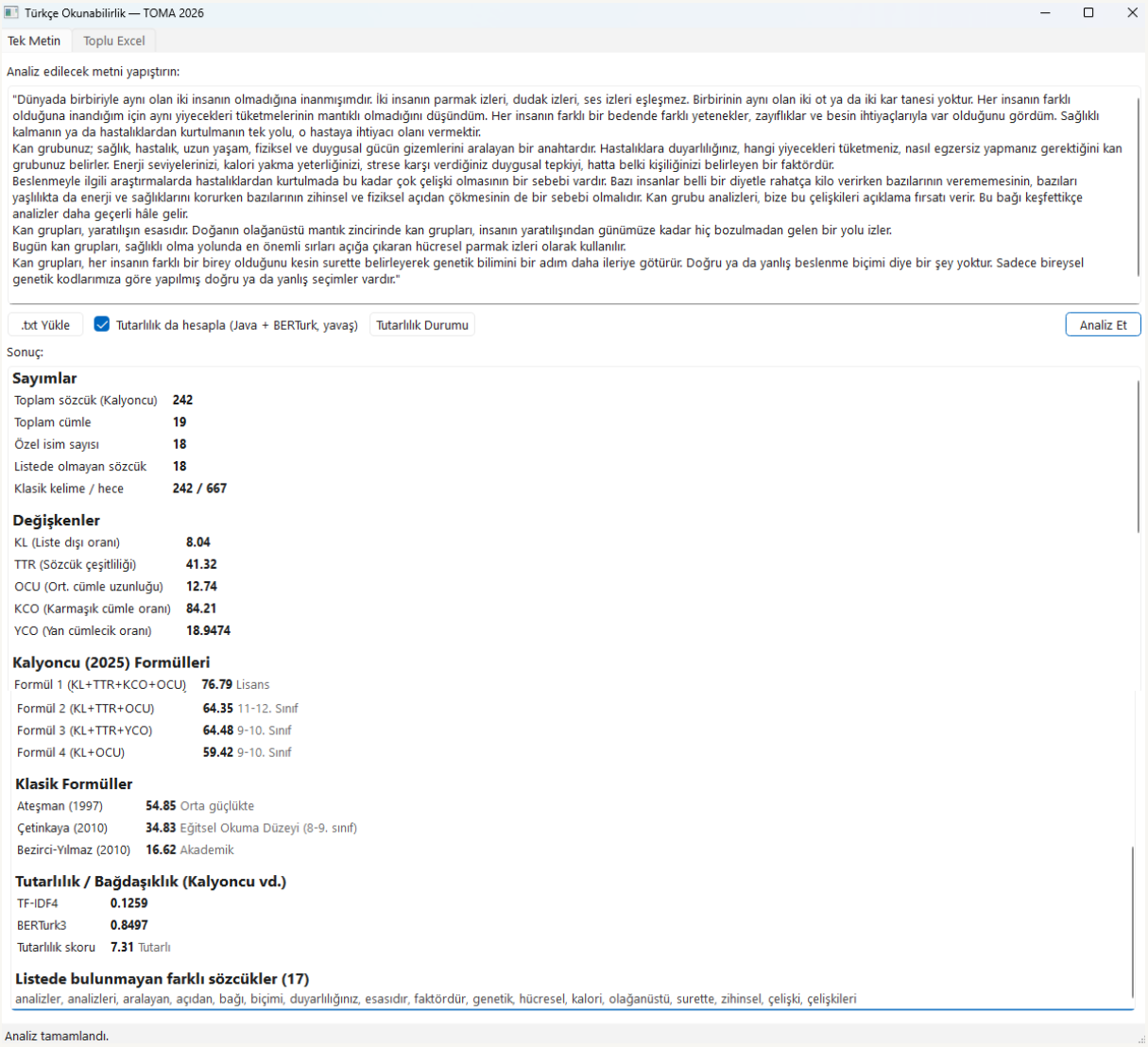
Bu çalışmada Türkçe metinler için otomatik metin analizi programı geliştirilmesi amaçlanmıştır. Bu amaç doğrultusunda, Claude Fable 5 aracılığıyla Python yazılımı ve çeşitli NLP araçları kullanılarak masaüstü bir program tasarlanmıştır. Tasarlanan programın hesaplama doğruluğu Pearson korelasyon analizi, intraclass korelasyon analizi, Bland-Altman analizi, normalize edilmiş mutlak hata analizi ve hesaplama sapması büyüklüğü analizi ile incelenmiştir. 57 metin üzerinde araştırmacıların hesaplamaları ile yazılım hesaplamaları kullanılarak yürütülen analizler sonucunda TOMA'nın Türkçe için geliştirilmiş okunabilirlik formüllerini ve bağdaşıklık/tutarlılık formülünü yüksek doğrulukla hesaplanabildiği saptanmıştır.

Formüllere yönelik hesaplamaların doğruluğu incelendiğinde, TOMA'nın Ateşman (1997), Çetinkaya (2010), Bezirci-Yılmaz (2010) formüllerini oldukça yüksek doğrulukla hesaplayabildiği görülmektedir. İlgili formüllerin mekanik yapısı göz önüne alındığında, bu durum esasen beklenen bir sonuçtur. Buna karşılık Kalyoncu (2025) tarafından geliştirilen K1 formülünde hata oranının görece yüksek olduğu (%5,77) ulaşılan sonuçlar arasındadır. İlgili formülün bünyesinde barındırdığı KCO değişkeninin anlamsal analizi gerekli kılan bir değişken olması ve bu nedenle hata oranının yüksek olması (%6,19) bu görece yüksek hatayı anlamlı hâle getirmektedir. Bu durumda ilgili parametrelerin ve ilgili parametreleri kullanan formüllere yönelik sonuçların yorumlanmasında temkinli olunması gerekmektedir. Kalyoncu (2025) tarafından üretilen diğer formüller olan K2, K3 ve K4'ün hata oranı ise %5'in altındadır. Normalize edilmiş mutlak hata oranlarının dışında korelasyon ve intraclass korelasyon analizlerinin sonuçları ele alındığında, ölçümlerin yüksek doğruluk gösterdiği sonucuna ulaşılmıştır. Bu sonuçlar bir arada değerlendirildiğinde, araştırmada geliştirilen yazılımdan elde edilen sonuçların güvenilir olduğu, yalnızca K1 sonuçlarının yorumlanmasında dikkatli olunması gerektiği ifade edilebilir. Bu doğrultuda K1 formülünün uygulanmasının gerekli görüldüğü durumlarda araştırmacıların kompleks cümle oranı değişkenini kendi hesaplamalarıyla teyit etmeleri yerinde olacaktır.

TOMA'da kullanılan bağdaşıklık/tutarlılık formülü ise Kalyoncu, Akdoğan, Memiş ve Kamışlı Öztürk (2026) tarafından gerçekleştirilen çalışmada geliştirilen ikinci formüldür. Esasında çalışmada geliştirilen ilk formülün kapsamı ve varyans açıklama oranı daha yüksektir. Ancak ilgili çalışmanın bünyesinde yer alan "zamir" değişkeni de yine anlamsal analiz gerektiren bir değişkendir. Bu nedenle mevcut araştırmada geliştirilen yazılıma mekanik olan ikinci formül de dâhil edilmiştir. Fakat geliştirilen yazılım formüllerle birlikte değişkenlerin değerlerini de sunmaktadır. Kullanıcılar, metinlerdeki zamir ve bağlaç oranlarını kendileri belirleyerek bağdaşıklık/tutarlılık konusunda daha kapsamlı sonuçlara ulaşabilirler.

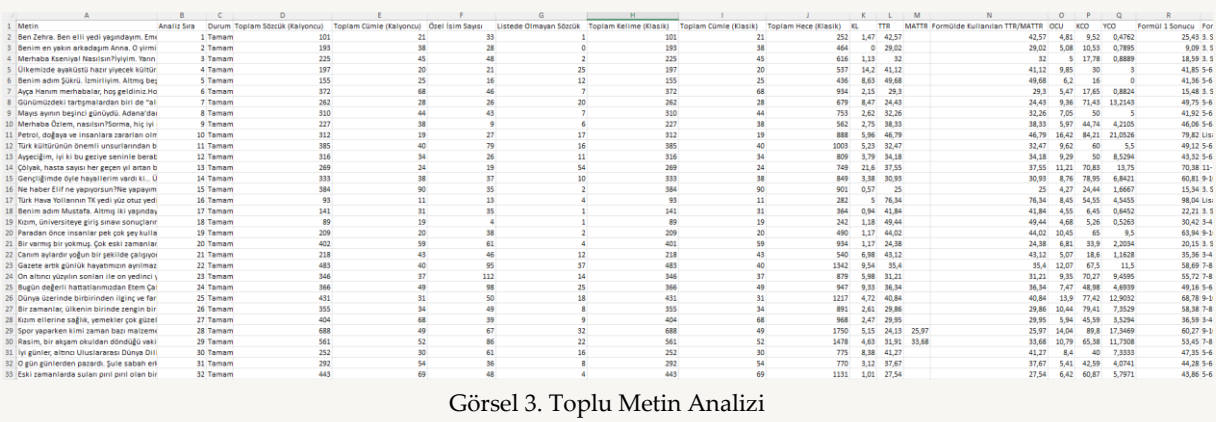
TOMA'nın hesaplama doğruluğunun yanı sıra araştırmacıların hangi formülleri kullanacağının da ele alınması gerekmektedir. Bilindiği üzere Türkçe için oluşturulmuş okunabilirlik formülleri birbirleri ile tutarlı sonuç vermemektedir (Kalyoncu ve Memiş, 2024). Aynı zamanda Kalyoncu (2025) tarafından üretilen farklı formüllerin hedef kullanım alanları belirli ölçüde farklılaşmaktadır. Literatürde gerçekleştirilen çalışmaların eğilimine bakıldığında ise en sık başvurulan formülün Ateşman (1997) formülü olduğu görülse (Memiş ve Kalyoncu, 2024) de yapılan incelemeler Çetinkaya (2010) formülünün Ateşman formülünden daha geçerli sonuçlar verdiğini göstermektedir (Kalyoncu ve Memiş, 2024). Ancak Türkçe için geliştirilmiş okunabilirlik formüllerinin bir arada karşılaştırıldığı güncel bir çalışma bulunmamaktadır. Bu doğrultuda gelecek çalışmalarda Türkçe için oluşturulmuş okunabilirlik formüllerini bir arada karşılaştıran ve güçlü kullanım alanlarını belirleyen çalışmalara ihtiyaç duyulduğu ifade edilebilir. Bunun yanı sıra 300-500 sözcükten oluşan metinlerde kompleks formüller olarak nitelendirilebilecek olan ve birçok değişkeni içeren K1, K2, K3 formüllerinin kullanılması, bu aralıktan daha kısa ya da uzun metinlerde K4 formülüyle birlikte Çetinkaya formülünün karşılaştırılmalı olarak kullanılması önerilmektedir. Ayrıca incelenen metinde sözcüksel karmaşıklık ön plana çıkıyorsa K4 formülünün daha geçerli sonuçlar vermesi, metin yüksek düzeyde sözcüksel karmaşıklık içermiyorsa Çetinkaya formülü ile K4 formülünün tutarlı sonuçlar vermesi beklenmektedir. Ancak metin uzunluğu 300-500 aralığının dışında kalan metinlerde K1, K2, K3 formüllerinden geçerli sonuç alınamaması olasıdır. Bağdaşıklık ve tutarlılık incelemelerinde ise araştırmacıların incelenen metindeki zamir ve bağlaç oranını tespit ederek birinci bağdaşıklık/tutarlılık formülünü uygulamaları daha geçerli sonuçlar alınmasına katkı sağlayacaktır.

Hesaplama doğruluğu ve formül kullanımının yanı sıra TOMA'nın kullanımına da değinilmesi mümkündür. Basit bir arayüze sahip TOMA'nın kullanımında esasen iki sekme bulunmaktadır. Birinci sekme doğrudan metinlerin yapıştırılarak analiz edilmesine imkân tanımaktadır. Aşağıda yer alan Görsel 2'de basit metin analizi gösterilmiştir.



Görsel 2. Basit Metin Analizi

Görselde görüldüğü üzere metnin ilgili bölüme yapıştırılması ve analiz edilmesiyle TOMA çıktılarına erişilmektedir. TOMA, kullanıcılarına değişkenleri ve formül sonuçlarını bir arada sunmaktadır. Bunun yanı sıra metinlerin nicel değerlerinin yanı sıra nitel kategorileri de çıktılarda belirtilmektedir. Örnek metin için araştırmacıların saptamaları K1=76.6, K2= 64.1, K3=63.3, K4=59.2, Ateşman=54.85, Çetinkaya=34.83 şeklindedir. Bu bağlamda somut örnek üzerinden de TOMA'nın hesaplamaları ile araştırmacı hesaplamalarını da karşılaştırmak mümkündür. TOMA'nın ikinci sekmesi ise toplu Excel analizine olanak tanıyan bölümdür. Bu sekmede metinlerin tek bir Excel dosyasına toplanarak toplu bir şekilde analiz edilmesi de mümkündür. Yalnız bu süreçte hazırlanan Excel dosyasında metinler sırasıyla tek hücreye yerleştirilmeli ve ilk hücrenin adı "Metin" olarak yapılandırılmalıdır. Bu şekilde gerçekleştirilen analizlere yönelik çıktı görüntüsü Görsel 3'te gösterilmiştir:



Görsel 3. Toplu Metin Analizi

Verilen bilgiler ve yapılan değerlendirmeler doğrultusunda alan yazınına bir metin analizi aracı kazandırıldığı söylenebilir. Nitekim alan yazında gerçekleştirilen okunabilirlik ve metin analizi çalışmalarında nispeten küçük metin örneklemleriyle çalışıldığı dikkat çekmektedir. Geliştirilen bu programla aynı zamanda büyük ölçekli metin analizleri ve derlem analizleri mümkün hâle gelmektedir. Araştırmacıların programı kullanarak zamandan tasarruf etmeleri ve inceleme kapsamlarını genişletebilmeleri beklenmektedir. Bununla birlikte TOMA'nın daha büyük örneklemler üzerinde bağımsız dış doğrulama yöntemleriyle ele alınması hesaplama doğruluğuna ilişkin daha kapsamlı kanıt sağlayabilir. Ayrıca konuyla ilgili gelişmelerin ve TOMA ile ilgili ayrıntıların "okunabilirlik.com.tr" adresinden takip edilmesi mümkündür. Tüm bu sonuçlardan hareketle geliştirilen bu programın okunabilirlik ve metin analizi çalışmalarında kullanılması, hâlihazırdaki okunabilirlik çalışmalarının kapsamının genişletilmesi önerilmektedir.

Sınırlılıklar

Araştırmada geliştirilen program, 57 metinden oluşan bir derlem ile analiz edilmiştir. Derlemdeki metinlerin değiştirilmesi bu çalışmada elde edilen değerlerin değişmesine neden olabilir. Diğer yandan programda yer verilen kompleks formüller %100 doğruluk oranıyla hesaplanamamaktadır. Ancak büyük bir derlemden elde edilen sonuçların hata oranlarının düşüklüğü dikkate alındığında TOMA'nın hesaplama doğruluğu yüksek bir araç olduğu ifade edilebilir. Buna ek olarak yazılım geliştirilme sürecindeki hesaplama kontrolleri ile nihai doğruluğun değerlendirilmesinde aynı derlemden yararlanılmıştır. Bu durumun araştırma sonuçlarını daha iyimser göstermiş olabilir. Gelecek çalışmalarda TOMA'nın doğruluğu bağımsız derlemler üzerinde sınanabilir.

Veri Erişilebilirliği: Çalışmada geliştirilen programa ve araştırmanın veri setine aşağıdaki bağlantıdan ulaşılabilir:

https://github.com/mrkalyoncu/TOMA-Turkce-Okunabilirlik-ve-Metin-Analizi-Programi-

Konuyla ilgili gelişmeler ve TOMA'nın detayları için https://okunabilirlik.com.tr/ bağlantısı ziyaret edilebilir.

Yazar Katkıları: Bu çalışmaya tüm yazarların katkı oranı %25'tir.

Çıkar Çatışması: Yazarlar arasında herhangi bir çıkar çatışması bulunmamaktadır.

Yapay Zekâ Kullanım Beyanı: Bu çalışmada tanıtılan araç olan TOMA, araştırmacıların yönlendirmeleri ve kontrolü doğrultusunda Claude Fable 5 ile geliştirilmiştir. Geliştirme sürecinin detayları yöntem bölümünde açıklanmıştır. Bununla birlikte çalışmanın genişletilmiş özetinin çevirisinde ChatGPT 5.6 aracından yararlanılmıştır.

Kaynakça

Aminovna, B. D. (2022). Importance of coherence and cohesion in writing. *Eurasian Research Bulletin*, *4*, 83–89.

Ateşman, E. (1997). Türkçede okunabilirliğin ölçülmesi. *Dil Dergisi*, *58*, 71–74.

Bezirci, B., & Yılmaz, A. E. (2010). Metinlerin okunabilirliğinin ölçülmesi üzerine bir yazılım kütüphanesi ve Türkçe için yeni bir okunabilirlik ölçütü. *Dokuz Eylül Üniversitesi Mühendislik Fakültesi Fen ve Mühendislik Dergisi*, *12*(3), 49–62. https://izlik.org/JA72ZC62RP

Budak, Y. (2005). Metinlerin okunabilirlik düzeyinin saptanmasına yönelik eleştirel bir bakış. *Eurasian Journal of Educational Research*, (21), 76–87.

Crossley, S. A., Kyle, K., & McNamara, D. S. (2016). The tool for the automatic analysis of text cohesion (TAACO): Automatic assessment of local, global, and text cohesion. *Behavior Research Methods*, *48*(4), 1227–1237. https://doi.org/10.3758/s13428-015-0651-7

Crossley, S. A., Kyle, K., & Dascalu, M. (2019). The tool for the automatic analysis of cohesion 2.0: Integrating semantic similarity and text overlap. *Behavior Research Methods*, *51*(1), 14–27. https://doi.org/10.3758/s13428-018-1142-4

Çetinkaya, G. (2010). *Türkçe metinlerin okunabilirlik düzeylerinin tanımlanması ve sınıflandırılması* [Doktora tezi, Ankara Üniversitesi]. Yükseköğretim Kurulu Ulusal Tez Merkezi.

Dascalu, M., Trausan-Matu, S., Dessus, P., & McNamara, D. S. (2015). Discourse cohesion: A signature of collaboration. In *Proceedings of the fifth international conference on learning analytics and knowledge* (pp. 350–354). Association for Computing Machinery. https://doi.org/10.1145/2723576.2723578

DuBay, W. H. (2004). *The principles of readability*. Impact Information.

Flesch, R. (1943). *Marks of readable style: A study in adult education*. Teachers College, Columbia University.

Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology, 32*(3), 221–233. https://doi.org/10.1037/h0057532

George, D., & Mallery, P. (2010). *SPSS for Windows step by step: A simple guide and reference, 17.0 update* (10th ed.). Allyn & Bacon.

Graesser, A.C., & McNamara, D.S. (2011), Computational analyses of multilevel discourse comprehension. *Topics in Cognitive Science*, *3*, 371-398. https://doi.org/10.1111/j.1756-8765.2010.01081.x

Graesser, A. C., McNamara, D. S., & Louwerse, M. M. (2003). What do readers need to learn in order to process coherence relations in narrative and expository text? In A. P. Sweet & C. E. Snow (Eds.), *Rethinking reading comprehension* (pp. 82–98). Guilford Press.

Graesser, A. C., McNamara, D. S., Louwerse, M. M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers, 36*(2), 193–202. https://doi.org/10.3758/BF03195564

Güneş, F. (2000). *Okuma-yazma öğretimi ve beyin teknolojisi*. Ocak Yayınları.

Gunning, R. (1952). *The technique of clear writing*. McGraw-Hill.

Hall, S., Basran, J., Paterson, K. B., Kowalski, R., Filik, R., & Maltby, J. (2014). Individual differences in the effectiveness of text cohesion for science text comprehension. *Learning and Individual Differences, 29*, 74–80. https://doi.org/10.1016/j.lindif.2013.10.014

Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. Longman.

Kalyoncu, M. R. (2025). *Türkçe için yeni bir okunabilirlik formülü* [Yüksek lisans tezi, Ondokuz Mayıs Üniversitesi]. Ondokuz Mayıs Üniversitesi Açık Erişim Sistemi.

Kalyoncu, M. R., Akdoğan, B. A., Memiş, M., & Kamışlı Öztürk, Z. (2026). Automatic measurement of cohesion and coherence in agglutinative languages: A cohesion and coherence formula for Turkish. *Glottometrics*, 61, 1-25.

Kalyoncu, M. R., & Memiş, M. (2024). Türkçe için oluşturulmuş okunabilirlik formüllerinin karşılaştırılması ve tutarlılık sorgusu. *Ana Dili Eğitimi Dergisi, 12*(2), 417–436. https://doi.org/10.16916/aded.1434650

Karatay, H. (2010). Bağdaşıklık araçlarını kullanma düzeyi ile tutarlı metin yazma arasındaki ilişki. *Mustafa Kemal Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 7*(13), 373–385.

Kincaid, J. P., Fishburne, R. P., Rogers, R. L., & Chissom, B. S. (1975). *Derivation of new readability formulas (Automated Readability Index, Fog Count, and Flesch Reading Ease Formula) for Navy enlisted personnel* (Research Branch Report 8–75). Chief of Naval Technical Training. https://doi.org/10.21236/ADA006655

Klare, G. R. (1984). Readability. In P. D. Pearson (Ed.), *Handbook of reading research* (pp. 681–744). Longman.

Leli, L. (2020). Analysis of coherence and cohesion on students' academic writing: A case study at the 3rd year students at English education program. *Alsuna: Journal of Arabic and English Language, 3*(2), 74–82. https://doi.org/10.31538/alsuna.v3i2.980

McLaughlin, G. H. (1969). SMOG grading: A new readability formula. *Journal of Reading, 12*(8), 639–646.

Memiş, M., & Kalyoncu, M. R. (2024). Türkçe ders kitaplarında okunabilirlik. In M. Memiş (Ed.), *Türkçe ders kitabı inceleme kılavuzu* (pp. 249–276). Nobel Akademik Yayıncılık.

Ozuru, Y., Dempsey, K., & McNamara, D. S. (2009). Prior knowledge, reading skill, and text cohesion in the comprehension of science texts. *Learning and Instruction, 19*(3), 228–242. https://doi.org/10.1016/j.learninstruc.2008.04.003

Smith, E. A., & Senter, R. J. (1967). *Automated readability index* (AMRL-TR-66-220). Aerospace Medical Research Laboratories.

Stenner, A. J., Burdick, H., Sanford, E. E., & Burdick, D. S. (2006). How accurate are Lexile text measures? *Journal of Applied Measurement, 7*(3), 307–322.

van Dijk, T. A. (1981). Discourse studies and education. *Applied Linguistics, 2*(1), 1–26. https://doi.org/10.1093/applin/II.1.1

Extended Abstract

Introduction

Readability refers to the ease or difficulty of reading arising from the structural features of a text, while the efficiency of the reading process is shaped by the alignment between the difficulty of the text and the reader's linguistic and cognitive characteristics (Kalyoncu, 2025). It is therefore important to determine whether different types of texts, ranging from educational materials to public documents, are appropriate for their intended audiences. Readability formulas are among the most widely used tools for determining text difficulty objectively and efficiently. Although several formulas have been developed for Turkish by Ateşman (1997), Çetinkaya (2010), Bezirci-Yılmaz (2010), and Kalyoncu (2025), their manual application, particularly to large text corpora, requires considerable time and effort. Moreover, the agglutinative structure of Turkish complicates the automated identification of variables such as lexical diversity, word familiarity, complex sentence ratio, and subordinate clause ratio. Although sophisticated tools such as Coh-Metrix (Graesser et al., 2004), TAACO (Crossley et al., 2016; Crossley et al., 2019), and Lexile Analyzer (Stenner et al., 2006) are available for English, the absence of an accessible software tool capable of jointly and automatically analyzing the readability and cohesion/coherence of Turkish texts represents an important gap.

This study aimed to address this gap by developing the Turkish Readability and Text Analysis Software (TOMA) and evaluating its computational accuracy. TOMA was designed to automate readability formulas developed for Turkish, generate a quantitative measure of cohesion/coherence, support both single-text and batch analyses, and thereby facilitate large-scale research.

Method

The study was conducted in two complementary stages: developing the software and validating its outputs against reference calculations. TOMA was developed as a Python 3.11-based desktop application with support from Claude Fable 5, under the researchers' subject-matter guidance and continuous supervision. The software architecture consists of two layers. The first is a Python package comprising independent modules that perform linguistic computations, while the second is a graphical user interface that supports both single-text analysis and batch analysis through Excel files. Various natural language processing tools, including Zeyrek, Stanza, Zemberek, BERTurk, and scikit-learn, were employed for morphological analysis, dependency parsing, lemmatization, sentence segmentation, semantic similarity, and lexical overlap. Before analysis, the texts were segmented into sentences and words, Turkish characters with circumflexes were normalized, and punctuation marks, numerals, and different word forms were processed according to the definition of each measure.

In addition to total word, sentence, and syllable counts, the software calculates 15 fundamental parameters, including mean sentence and word length, word-list score, type-token ratio (TTR), moving-average type-token ratio (MATTR), complex sentence ratio, subordinate clause ratio, TF-IDF, and BERTurk similarity. Using these parameters, TOMA computes eight formulas: the Ateşman, Çetinkaya, and Bezirci-Yılmaz formulas; Kalyoncu's K1, K2, K3, and K4 formulas; and a cohesion/coherence formula. The second cohesion/coherence formula, which is based on TF-IDF and BERTurk (Kalyoncu et al., 2026), was incorporated into the software. The pronoun variable included in the more comprehensive first formula was excluded from automated computation because demonstrative pronouns cannot be fully distinguished from demonstrative adjectives through purely rule-based procedures.

The computational accuracy of the software was evaluated using a heterogeneous corpus of 57 texts comprising 17,678 words. Constructed through extreme case sampling, the corpus included texts representing expository, narrative, dialogic, autobiographical, academic, historical, and spoken discourse. Text length ranged from 89 to 686 words, with a mean length of 310.14 words. Before the software was developed, a subject-matter expert calculated all parameter and formula values for the texts, and these values were treated as reference measurements. The relationship and absolute agreement between the final software outputs and the reference values were examined using Pearson correlation analysis, intraclass correlation coefficients (ICCs) based on a two-way mixed-effects model, normalized mean absolute error (NMAE), and Bland–Altman analyses.

Results

The findings showed that TOMA generally produced results that closely matched the reference calculations, demonstrating a very high level of computational accuracy. Pearson correlation coefficients ranged from .952 to 1.000 for the parameters and from .964 to 1.000 for the eight formulas. Correlations reached 1.000 for total syllable count, TF-IDF, BERTurk, and cohesion/coherence calculations. At the formula level, correlations of .997 for Ateşman, .998 for Çetinkaya, .999 for Bezirci-Yılmaz, and 1.000 for cohesion/coherence were obtained. ICC values ranging from .936 to 1.000 demonstrated not only strong linear relationships but also very high absolute agreement. NMAE remained below 5% for 17 of the 19 measures and fell below 1% for word and syllable counts, TF-IDF, BERTurk, Ateşman, Çetinkaya, and cohesion/coherence. The highest error rates were observed for the complex sentence ratio (6.19%) and the K1 formula, which incorporates this variable (5.77%). Nevertheless, the ICC values for both measures remained above .95. The Bland–Altman results confirmed that systematic bias was low for most variables and formulas, while wider limits of agreement were observed for the complex sentence ratio and K1. TOMA also tended to produce slightly higher scores for K1–K4 and slightly lower scores for the Ateşman and Çetinkaya formulas.

Conclusion

The findings indicate that TOMA is a reliable and functional tool for automatically analyzing the readability and cohesion/coherence characteristics of Turkish texts. However, high computational accuracy does not imply that all formulas included in the software are equally valid for every text type and length.

The software offers a simple analysis interface into which a single text can be pasted, as well as a batch-analysis interface that operates through Excel files. These features enable researchers to analyze both individual texts and large corpora within a short period. TOMA presents variable values, formula scores, and their corresponding qualitative categories together, thereby facilitating the verification of the results. When K1 is used, however, the complex sentence ratio should also be independently verified by the researcher. For texts between 300 and 500 words in length, the use of K1, K2, and K3 is recommended; for texts outside this range, K4 and the Çetinkaya formula should be used comparatively. Users seeking a more comprehensive assessment of cohesion/coherence may additionally determine pronoun and conjunction ratios and apply the first formula.

The principal limitation of the study is that the validation was conducted using 57 texts and that the same corpus was used both for the checks performed during software development and for the final evaluation. This may have resulted in somewhat optimistic estimates of computational accuracy. TOMA should therefore be externally validated using larger, independent corpora comprising texts of different types, lengths, and periods. Despite this limitation, the low error rates and high levels of absolute agreement provide strong evidence that TOMA can be used reliably in readability and text-analysis research. By reducing researchers' computational burden, the openly accessible software is expected to facilitate studies with larger samples and broaden the scope of Turkish readability research.