

Dil Eğitimi ve Araştırmaları Dergisi
(DEvAD)

Journal of Language Education and Research
(JLERE)

ISSN: 2149-5602 | e-ISSN: 2149-5602

2026 | 12 (1) | 377-397 |

DOI: 10.31464/jlere.1858803

Makale Türü: Araştırma Makalesi

Önerilen Atıf

Şimşek, P. H. ve Özdemir, S. (2026). Yabancı dil olarak Türkçe öğrencilerinin yazdıkları metinlerde ve ChatGPT tarafından yazılan metinlerde karmaşıklık, doğruluk, akıcılık ve işlevsel yeterlik. *Dil Eğitimi ve Araştırmaları Dergisi*, 12 (1), 377–397.

Makale Geçmişi

Geliş: 07.01.2026

Düzeltilme: 28.03.2026

Kabul: 03.04.2026

Yayın: 30.04.2026

Yazar Katkı Oranı (CRediT)

Yazar 1: Veri toplama, analiz, yazım - özgün taslak

Yazar 2: Kavrasallaştırma, yöntem, analiz, yazım - inceleme ve düzenleme

Etik Açıklama

Bu araştırma Bartın Üniversitesi Sosyal ve Beşeri Bilimler Etik Kurulu tarafından 16.11.2023 tarih ve 17 sayılı karar ile onaylanmıştır.

Yapay Zekâ Kullanımı

Bu çalışmanın hazırlanmasında üretken yapay zekâ katkısı yoktur.

Yabancı Dil Olarak Türkçe Öğrencilerinin Yazdıkları Metinlerde ve ChatGPT Tarafından Yazılan Metinlerde Karmaşıklık, Doğruluk, Akıcılık ve İşlevsel Yeterlik

Complexity, Accuracy, Fluency, and Functional Adequacy in Texts Written by Learners of Turkish as a Foreign Language and Texts Generated by ChatGPT

Pınar Havva Şimşek^{1*}, Serpil Özdemir²

¹ Bartın Üniversitesi, Yabancılar Türkçe Öğretimi Anabilim Dalı Yüksek Lisans Öğrencisi, Bartın, Türkiye—
ORCID: /0009-0002-7020-0721

² Bartın Üniversitesi, Türkçe ve Sosyal Bilimler Eğitimi Bölümü, Bartın, Türkiye — ORCID: 0000-0002-8063-8690

Sorumlu yazar: serpilozdemir@bartin.edu.tr.

Öz

Bu araştırmanın amacı yabancı dil olarak Türkçe öğrencilerinin yazdıkları metinler ile ChatGPT tarafından oluşturulan metinleri karmaşıklık, doğruluk, akıcılık ve işlevsel yeterlik açısından karşılaştırmaktır. Çalışma karşılaştırmalı betimsel araştırma deseni ile yürütülmüştür. Araştırmada A2 ve B2 düzeyindeki öğrencilerin, altı farklı konuda yazdıkları yetmiş iki metin ve aynı konularda ChatGPT'ye yazdırılan altı metin incelenmiştir. Metinlere nicel içerik analizi uygulanmıştır. Araştırma sonucunda ChatGPT'nin daha yüksek karmaşık cümle oranlarına sahip olduğu, ancak düzeylere göre tutarlılık göstermediği belirlenmiştir. ChatGPT, cümle ve sözcük düzeyinde öğrencilerden daha doğru metinler üretmiştir. Farklı sözcük kullanımı açısından ChatGPT'nin ortalamaları daha yüksek bulunmuş, ancak anlamlı sözcük oranı bakımından öğrenciler daha yüksek değerlere ulaşmıştır. İşlevsel yeterlik açısından ChatGPT daha yüksek puanlar almıştır. Ayrıca A2 düzeyindeki bazı öğrenci metinlerinin B2 düzeyindekilere göre daha nitelikli olduğu da görülmüştür. Sonuçlar, ChatGPT'nin cümle ve sözcük doğruluğunu geliştirmede ve işlevsel yeterliğin sağlanmasında yabancı dil olarak Türkçe öğrenenlere potansiyel bir destek aracı olabileceğini göstermektedir. ChatGPT'den metni planlama veya yazılan metni gözden geçirme aşamalarında yararlanılması önerilmektedir.

Anahtar Sözcükler: akıcılık, ChatGPT, doğruluk, işlevsel yeterlik, karmaşıklık, yabancı dil olarak Türkçe öğretimi

Abstract

The purpose of this study is to compare texts written by learners of Turkish as a foreign language with texts generated by ChatGPT in terms of complexity, accuracy, fluency, and functional adequacy. The study was conducted using a comparative descriptive research design. The study examined 72 texts written by learners at the A2 and B2 levels on six different topics and 6 texts written by ChatGPT on the same topics. Quantitative content analysis was applied to the texts. The results of the study determined that ChatGPT had a higher percentage of complex sentences but did not show consistency across levels. ChatGPT produced more accurate texts than the learners at the sentence and word levels. ChatGPT had higher averages in terms of different word usage, but the learners achieved higher values in terms of meaningful word ratio. In terms of functional adequacy, ChatGPT scored higher. It was also observed that some learner texts at the A2 level were more qualified than those at the B2 level. The results indicate that ChatGPT has the potential to serve as a support tool for learners of Turkish as a foreign language in improving sentence and word accuracy and ensuring functional adequacy. It is recommended to use ChatGPT during the text planning or review stages.

Keywords: fluency, ChatGPT, accuracy, functional adequacy, complexity, teaching Turkish as a foreign language

1. Giriş

Yabancı dilde yazma becerisi, dilsel yeterliklerin gelişimini somut olarak gösteren, üretime dayalı bir dil becerisidir. Yazılı üretim sürecinde dil bilgisi kurallarına uygun cümleler kurmak anlamın iyi yapılandırılmasına destek olur, ancak anlamlı, tutarlı ve türe uygun metinler üretmeyi garantilemez. Bu

nedenle yazma becerisi hedef dile özgü yapıları doğru kullanmanın yanında düşünceleri organize etme ve farklı türde metinlerle etkili yazılı iletişim kurma yeterliği gerektirir (Richards ve Renandya, 2002). İletişimin etkili olabilmesinde sosyo-kültürel unsurların da bilinmesi önemlidir. Çünkü yabancı dilde öğrenciler yetkinleştikçe yazılarında muhatabına göre üslup kullanmaları beklenir (Council of Europe, 2020).

Dilsel, bilişsel, kültürel unsurları yetkin biçimde kullanmayı gerektirdiği için yabancı dilde yazma becerisi zor gelişen bir beceridir. Türkçeyi yabancı dil olarak öğrenen bireylerin yazılı anlatımları incelendiğinde dil bilgisel hatalar, sözcük seçimi sorunları, bağlaç ve zaman uyumsuzlukları, cümle kuruluş hataları ve paragraf bütünlüğü eksiklikleri, tutarlılık gibi sorunlar görülmektedir (Baş ve Turhan, 2017; Bülbül ve Karakaş, 2024; Özdemir ve Şimşek, 2025). Bu hataların temelinde, öğrencilerin hedef dilde yeterince yazılı girdiyle karşılaşmamış olmaları, ana dillerinden yaptıkları doğrudan aktarımlar ve yazma sürecine ilişkin stratejik yetersizlikler yer almaktadır. Yazma becerisinin gelişimi, bireysel farkındalık ve etkili geri bildirim mekanizmalarıyla doğrudan ilişkilidir, ancak geleneksel öğretim ortamlarında bireysel yazma hatalarının belirlenmesi ve anında geri bildirim sunulması her zaman mümkün olamamaktadır. Bu durumda, üretici yapay zekâ araçları insan benzeri metin oluşturma yetenekleriyle dikkat çekmiş, dil öğrenme ve öğretme süreçlerinde potansiyel bir yardımcı araç konumuna gelmiştir (Zhai, 2022; Zhang ve Tur, 2024). Yapay zekâ sistemleri eğitildikleri verilerle insan benzeri öğrenme, karar verme ve problem çözme görevlerini gerçekleştirebilmektedirler (Bellman, 1978; Heller, 2019). Çeşitli türlere ve dillere yayılan büyük metin veri kümeleri üzerinden eğitilen bu araçlar, bağlamsal olarak tutarlı ve dil bilgisi açısından doğru metinler üretebilirler. Dil kalıplarının, sözdiziminin ve semantiğin inceliklerini öğrenebilirler. Böylece, bağlama uygun olan ve insan tarafından yazılmış içerikten ayırt edilemeyen metinler üretebilirler (Koşaner vd., 2023, s.146).

ChatGPT ve Yazılı Metinler

OpenAI tarafından geliştirilen ChatGPT, insan metnine benzer dil işleme modelleri oluşturmak için derin öğrenme tekniklerini kullanmaktadır. Büyük miktarda metin verisi üzerinde eğitilmiş olan bu sohbet robotu, doğal bir şekilde dil girdilerini anlayarak bağlamsal yanıtlar üretebilmekte, metin özetleme, diller arası çeviri yapma, insan metnine benzer geri bildirimler verme, içerik oluşturma gibi görevleri yerine getirebilmektedir (Öztürk Demir ve Çetin, 2025). ChatGPT bu özellikleriyle yabancı dil öğretiminde hızlı bir şekilde yer almıştır. Yabancı dil olarak İngilizce öğrenenlerle yapılan araştırmalarda öğrencilerin ChatGPT aracılığıyla yazma becerisini geliştirmeye yönelik çalışmalar yapılmıştır. Bu çalışmaların sonucunda dil bilgisel doğruluğun arttığı, kelime dağarcığının geliştiği, paragraf tutarlılığının arttığı, türe uygun yazma becerisinin geliştiği görülmüştür (An, 2024; Song ve Song, 2023).

ChatGPT'nin doğruluk, akıcılık, karmaşıklık ve tutarlılık gibi dilsel nitelikler açısından nasıl metinler ürettiği ve bu metinlerin yabancı dil öğrenenlerin metinleriyle karşılaştırılması genellikle İngilizcenin yabancı dil olarak öğretiminde ve farklı kültürlerde araştırılmıştır. Bu araştırmalarda bazı durumlarda ChatGPT üstünlüğü görülürken bazı durumlarda öğrenci üstünlüğü söz konusudur. Fredrick ve Craven (2025) tarafından Arap öğrencilerle yapılan araştırmada ChatGPT'nin daha fazla sözcüksel çeşitliliğe ve sözdizimsel karmaşıklığa sahip metinler üretebildiği ancak iletişimsel uygunlukta yetersiz kaldığı görülmüştür. Zhou vd. (2023) tarafından Çinli öğrencilerle yapılan araştırmada ise ChatGPT'nin anlatı oluşturmada, somut sözcüklerle kolayca anlamayı sağlamada ve gönderimleri tutarlı bir şekilde kullanmada öğrencilerden daha başarılı olduğu, ancak derin anlamı oluşturan nedensel ve amaçsal ilişkileri kurmada öğrencilerin gerisinde kaldığı belirlenmiştir. Herbold vd. (2023) Alman öğrencilerin yazdıkları tartışmacı metinler ile ChatGPT 3.5 ve ChatGPT 4 sürümlerini karşılaştırmışlardır. Araştırmada ele alınan

karmaşıklık, sözcüksel çeşitlilik, işlevsel yeterlik ölçütlerinde sırasıyla öğrenciler, ChatGPT 3.5 ve ChatGPT 4 başarılı bulunmuştur.

ChatGPT'nin Türkçe dil yeterliğini insan üretimi metinlerle dilbilimsel açıdan karşılaştıran tek araştırma Yıldız Durak vd. (2025) tarafından ana dili öğrencileriyle yapılmıştır. Bu çalışmada ChatGPT, Gemini ve Bing tarafından üretilen metinlerle Öğretim Teknolojileri dersini alan eğitim fakültesi öğrencilerinin yazdıkları metinler karşılaştırılmıştır. "Eğitim 4.0, endüstri 4.0 ve toplum 5.0 kavramlarını nasıl ilişkilendirirsiniz?" konusunda yazılan tartışmacı metinler cümle uzunluğu ve akıcılık açılarından analiz edilmiştir. Araştırma sonucunda ChatGPT metinlerinde cümle uzunluğunun, görevli sözcüklerin daha fazla olduğu, öğrenci metinlerinde ise farklı sözcüklerin daha fazla olduğu belirlenmiştir.

Metinlerde Karmaşıklık Doğruluk, Akıcılık ve İşlevsel Yeterlik

İkinci dil ediniminde yazma performansının değerlendirilmesinde karmaşıklık, akıcılık ve doğruluk kavramları kritik bir öneme sahiptir (Larsen-Freeman, 2006). Larsen-Freeman'ın (1978) dil gelişimini ölçmeye yönelik geliştirdiği indeksler, bu üç boyutun birbirleriyle olan ilişkisini ve öğrenci gelişimini nasıl yansıttığını gösteren temel çalışmalardan biridir. Bu kavramlar, sadece öğrenci metinlerinin değil, aynı zamanda yapay zekâ tarafından üretilen metinlerin de analizi için geçerli bir çerçeve sunmaktadır. Bununla birlikte karmaşıklık, akıcılık ve doğruluk çerçevesinde yapılan incelemeler metnin dil bilgisi açısından doğruluğunu ortaya koymaktadır, ancak metinlerin işlevsel yeterliğini belirleyememektedir. Bu nedenle işlevsel yeterlik ölçütünün de dikkate alınması önerilmektedir (Zhang ve Tur, 2024). Söz konusu ölçütlerin birlikte ele alınması metinlerin hem dilsel açıdan doğruluğunu hem de anlamsal açıdan yeterliğini incelenmeyi sağlamaktadır.

Karmaşıklık, metinlerin yapısal zenginliğini, cümle uzunluğunu, yan cümle kullanımını ve bağlama uygun yapılarla çeşitlenmesini ifade eder. Yazının üst düzey bir üretim göstergesi olan karmaşık yapıların kullanımı (Kuiken ve Vedder, 2017), ikinci dilde yazma becerisinin gelişmişliğinin bir göstergesidir. İngilizcenin karmaşık cümle yapısının belirlenmesinde ortalama cümle uzunluğu, yan cümle oranı ve t-birimler dikkate alınmaktadır. T-birimi, ana cümle ve ona bağlı veya gömülü tüm yan cümleler ve cümle dışı yapılardan oluşmaktadır (Hunt, 1965'ten akt. Kuiken, 2023).

Türkçede yapısına göre cümle sınıflaması konusunda uzmanların farklı sınıflamalar yaptıkları görülmektedir. Basit cümlelerin tek bir yüklemi olduğu konusunda uzlaşa sağlanmıştır. Ancak sıfat-fiil ve zarf-fiillerin bulunduğu cümleleri basit cümle kabul eden görüşler olduğu gibi (Ergin, 1985; Karahan, 2000) girişik birleşik cümle kabul eden görüşler de mevcuttur (Dizdaroğlu, 1976; Gencan, 1979). Bununla birlikte dil bilgisi öğretiminde yerleşmiş genel kabule göre yapılarına göre cümleler basit, birleşik (girişik birleşik, iç içe birleşik, şartlı birleşik, ki'li birleşik), sıralı (bağımsız sıralı, bağımlı sıralı) ve bağlı cümleler olarak ele almaktadır (Balyemez, 2021).

Doğruluk, metnin sözcük ve cümle düzeyinde hatasız olması ile ilgilidir. Doğruluk, dili yapı ve anlam bakımından uygun şekilde kullanmayı gerektirmektedir. Sözcüğün doğru yazılması, anlamca uygun sözcüğün seçilmesi, söz diziminin doğru yapılandırılması ve mantıklı olması doğruluğu sağlayan ölçütlerdir. Hataların yüksek oranda olması metnin anlaşılmasını etkileyen bir unsurdur. Yabancı dil öğrencilerinin tam bir doğruluk ile dili kullanmaları beklenemez. Doğruluk oranının ne kadar olması gerektiği konusunda ise kesin bir bilgi yoktur. Celce-Murcia (1991) 100 kelimedeki ortalama 7.2 sözcüğün hatalı olmasının akademik yazmada kabul edilemez olduğunu belirtmektedir.

Akıcılık, metnin anlam bütünlüğünün sürdürülmesi, ifadelerin doğal ve tutarlı olması ile yazma sürecinde yazarın düşüncelerini kesintisiz şekilde aktarabilmesidir (Larsen-Freeman, 2006). Akıcılık, hem

üretim sürecindeki hız hem de metnin içsel bütünlüğüyle doğrudan ilgilidir. Akıcılık, düşüncenin yazılı olarak süreklilik içinde aktarılması, gereksiz tekrarların ve duraksamaların olmaması ile ilişkilidir.

Akıcılığın temel ölçütü olarak genellikle metindeki toplam sözcük sayısı ele alınmaktadır (Storch, 2005). Öğrencilerin sözcük dağarcığının genişliği ve sözcükleri doğru bağlamda kullanabilme yeteneği, metnin anlamsal niteliğine katkı sağlayan bir unsurdur. Ancak söz varlığının etkili kullanılmasında sadece sözcük sayısı yeterli değildir, aynı sözcüklerin sürekli tekrar edildiği bir metin akıcılıktan uzaklaşacaktır. Bu nedenle söz varlığının etkili kullanılmasında farklı sözcük kullanımı akıcılığın bir ölçütü durumundadır. Akıcılığı belirleyen bir ölçüt de anlamlı sözcük kullanımınıdır. Anlamlı sözcükler tek başına anlam taşıyan sözcüklerdir. Görevli sözcükler ise tek başına anlamı olmayan, cümledeki diğer sözcüklerin yardımı ile anlam kazanan, anlamlı sözcüklerin cümlede kullanılmasını sağlayan sözcüklerdir (Ergin, 1985; Korkmaz, 2009). Bu sözcüklerin işlevsel kullanımı, metinlerde anlamsal kopuklukların oluşmasını engeller.

Türkçede uzmanların yaptıkları anlamlı ve görevli sözcük sınıflandırmaları incelendiğinde Ergin'in (1985), sözcükleri anlam ve görevleri bakımından üç ana gruba ayırdığı görülür. Bunlar isimler, fiiller ve edatlardır. Ergin, görevli sözcükleri edat başlığı altında inceler. Edatları ünlem edatları, bağlama edatları ve son çekim edatları olarak ele alır. Korkmaz (2009), isimler ve fiilleri anlamlı sözcükler olarak ele alır. İsim soylu sözcükler kapsamında isimler, sıfatlar, zamirler, zarflar yer almaktadır. Edatlar ve bağlaçları görevli sözcükler kapsamında değerlendirir, ünlemleri ise görevli ve anlamlı sözcükler olarak değerlendirir. Korkmaz, ünlemlerin yalnız başına anlamlı olduğu, bir ad gibi çekimlendiği veya türetildiği durumlarda anlamlı sözcük işlevi görmesi nedeniyle anlamlı sözcük sayılabileceğini belirtmektedir. Çekime girmeyen, doğrudan bir kavrama karşılık gelmeyen ünlemleri ise görevli sözcük olarak değerlendirmektedir.

İşlevsel yeterlik, bir metnin iletişim amacına hizmet edip etmediğini ve yazı türüne uygunluk gösterip göstermediğini sorgular. Kuiken ve Vedder (2017), işlevsel yeterliği; amacın gerçekleştirilmesi, içeriğin tutarlılığı ve hedef okuyucuya uygunluk olarak açıklar. Bu yeterlik, yazılı bir görevin, belirlenen amaç ve beklentilere ne kadar uygun şekilde tamamlandığının bir göstergesidir. İşlevsel yeterlik ölçütleri "içerik, görev gereksinimleri, anlaşılabilirlik, tutarlılık ve bağdaşıklık" olmak üzere dört temel boyuttan oluşur. Bu boyutları değerlendirme ölçütleri şunlardır:

İçerik ölçütü, yazılı metinde sunulan bilgi birimlerinin sayısının yeterli olup olmadığını ve konuyla alakasını değerlendirir. Fikirlerin birbiriyle tutarlılığı ve metnin genel zenginliği bu boyutun temelini oluşturur.

Görev gereksinimleri ölçütü, öğrencinin yazma görevinin tür, söz eylemleri ve hitap tarzı gibi gerekliliklerini ne kadar başarılı bir şekilde yerine getirdiğini inceler. Görevin amacına tam olarak ulaşıp ulaşılamadığı bu boyutta sorgulanır.

Anlaşılabilirlik ölçütü, metnin amacını ve içerdiği fikirleri anlamak için ne kadar çaba gerektiğine odaklanır. Metnin ne kadar açık, net ve okuyucu tarafından kolayca anlaşılabilir olduğu bu boyutta değerlendirilir.

Tutarlılık ve bağdaşıklık ölçütü, metnin mantıksal bir bütünlük içinde olup olmadığını, fikirler arasındaki geçişlerin ve bağlaçların kullanımının yeterliliğini dikkate alır.

Kuiken ve Vedder (2017) tarafından belirlenen bu boyutlar hem yazılı anlatımdaki işlevsel yeterliği kapsamlı bir şekilde değerlendirmek için bir çerçeve sunmaktadır hem de yazma öğretiminin hedefleriyle uyumludur.

Araştırmanın Amaçları ve Araştırma Sorusu

Yazma becerisi; dil bilgisi, sözcük bilgisi, metin türü bilgisine dayanan yeterlikler gerektirmektedir. Dolayısıyla dil öğrenenler için zorlayıcı bir üretim becerisidir. Öğrenci metinlerinin hata analizine dayalı araştırmalarda dil bilgisi, sözcük seçimi, söz dizimi (Özdemir ve Şimşek, 2025) ve tutarlılık (Baş ve Turhan, 2017; Bülbül ve Karakaş, 2024) sorunları görülmektedir. Yapay zekâ araçları dil öğrenenler için potansiyel bir destek aracı olarak görülmektedir. Ancak bu araçlar eğitildikleri veri setine bağlı sonuçlar verdikleri için her zaman güvenilir olmayabilirler. Bu nedenle eğitimcilerin yapay zekâyı derslerinde kullanmadan önce kendi dillerinde ve konu alanlarında yapay zekânın yeterlik durumunu incelemeleri gerekmektedir. İngilizce için öğrenci ve ChatGPT metinlerinin niteliklerini karşılaştıran araştırmalar farklı kültürlerden örnekler üzerinde araştırılmaktadır (Fredrick ve Craven, 2025; Herbold vd. 2023; Zhou vd. 2023;). Türkçe için yapılan tek çalışma ana dili öğrencilerinin metinleri üzerinden yürütülmüştür (Yıldız Durak vd., 2025). Söz konusu araştırmanın bulguları akıcılık ile sınırlıdır. Ana dili kullanıcıları ile yabancı dil kullanıcılarının dilsel yeterlikleri farklı olduğundan farklı araştırmalarla konu ele alınmalıdır. Yabancı dil olarak Türkçe öğrencilerinin metinleri ile ChatGPT metinlerini dilbilimsel açıdan karşılaştıran bir araştırmaya rastlanmamıştır. Bu araştırma literatürdeki bu boşluktan hareketle yapılmıştır. Araştırmada ChatGPT'nin Türkçe dil bilgisi, söz varlığı, cümle yapısı ve işlevsel yeterlik açısından doğru metinler üretip üretmediğinin incelenmesi, sınırlılıklarının belirlenmesi, öğrenci metinleri ile karşılaştırılması amaçlanmaktadır. Çalışma sorusu “Yabancı dil olarak Türkçe öğrencilerinin metinleri ile ChatGPT tarafından oluşturulan metinlerin karmaşıklık, doğruluk, akıcılık ve işlevsel yeterlik açısından görünümü nasıldır?” şeklinde belirlenmiştir.

2. Yöntem

2.1. Araştırma Modeli / Deseni

Bu araştırma, karşılaştırmalı betimsel araştırma deseni çerçevesinde yürütülmüştür. Karşılaştırmalı betimsel araştırmalar, mevcut durumları herhangi bir müdahalede bulunmaksızın belirli değişkenler açısından karşılaştırmayı amaçlayan, deneysel olmayan araştırma türleridir (Cantrell, 2011). Bu tür araştırmalarda araştırmacı, var olan iki ya da daha fazla grubun belirli ölçütler doğrultusunda özelliklerini betimleyerek aralarındaki benzerlik ve farklılıkları ortaya koyar. Bu çalışmada yabancı dil olarak Türkçe öğrenen A2 ve B2 düzeyindeki öğrencilerin yazdıkları metinler ile aynı konularda ChatGPT tarafından üretilen metinler; karmaşıklık, doğruluk, akıcılık ve işlevsel yeterlik ölçütleri açısından karşılaştırılmıştır. Araştırmada herhangi bir öğretimsel müdahale gerçekleştirilmemiş ve mevcut metinler belirlenen ölçütlere göre incelenerek betimsel karşılaştırmalar yapılmıştır.

Metinlerin çözümlemesinde nicel içerik analizi yaklaşımından yararlanılmıştır. İçerik analizi, metinlerdeki belirli kategorilerin sistematik biçimde tanımlanması, kodlanması ve sayısallaştırılması yoluyla örüntülerin ortaya konmasını sağlayan bir yöntemdir (Krippendorff, 2018). Bu araştırmada karmaşıklık, doğruluk ve akıcılık göstergeleri önceden tanımlanmış kategoriler doğrultusunda sayısallaştırılmış; frekans ve yüzdelik değerler hesaplanarak raporlanmıştır. İşlevsel yeterlik ise dereceli puanlama anahtarı kullanılarak değerlendirilmiştir. Dolayısıyla çalışma, deneysel bir müdahale içermeyen, mevcut performans ürünlerinin önceden belirlenmiş ölçütler doğrultusunda karşılaştırıldığı karşılaştırmalı betimsel bir araştırma niteliği taşımaktadır.

2.2. Çalışma Kümesi

Araştırma temel ve orta düzeyde Türkçe öğrenen öğrencilerle sınırlandırılmıştır. C1 düzeyindeki öğrencilerin metinlerinde söz varlığı birikiminin fazla olması, karmaşık dil yapılarını kullanarak farklı türlerde tutarlı metinler üretebilme yeterliklerinin gelişmiş olması nedenleriyle C1 düzeyindeki öğrenci metinleri ile ChatGPT metinlerini karşılaştırmanın anlamlı bir sonuç vermeyebileceği öngörülmüştür. Bu nedenle temel ve orta düzeyleri temsilen A2 ve B2 düzeyindeki öğrencilerle araştırma yürütülmüştür. Araştırma Bartın Üniversitesi Türkçe Öğretim Merkezinde Türkçe öğrenen öğrencilerle yürütülmüştür. Araştırmacıların aynı kurumda görev yapması nedeniyle araştırmanın örnekleme kolay ulaşılabilir örneklem özelliği taşımaktadır. Araştırmaya katılanlar Çizelge 1’de sunulmuştur.

Çizelge 1. Katılımcıların dil düzeylerine ve ülkelere göre dağılımı

Düzyer	Kazakistan	Endonezya	Yemen	Mısır	Azerbaycan	Gabon	Fas	Rusya	Kongo	Mali	Zambiya	Demokratik Kongo Cum.	Ekvator Ginesi	Filistin	Tacikistan	Toplam
A2	12	2	2	1	1	0	1	1	1	0	0	0	0	0	0	21
B2	3	8	1	1	1	2	0	0	0	1	1	1	1	1	1	22
Toplam	15	10	3	2	2	2	1	1	1	1	1	1	1	1	1	43

Çizelge 1’de görüldüğü gibi araştırmaya 15 ülkeden toplam 43 öğrenci katılmıştır. A2 düzeyinde öğrencilerin yarısından fazlasının Kazakistanlı olduğu (%57), B2 düzeyinde ise Endonezyalı öğrencilerin diğer ülkelere göre daha yoğun olduğu görülmektedir (%36). Araştırmanın yürütüldüğü kurumda A2 düzeyinde 30, B2 düzeyinde 27 öğrenci derslere kayıtlıdır. Ancak çalışmanın yapıldığı günlerde derse katılmayan öğrencilerin olması veya öğrencinin çalışma için gönüllü olmaması nedenleriyle yazma konuları her defasında aynı öğrenciler tarafından yazılmamıştır. Her konuda minimum 12 metin yazılmıştır. 12 metinden fazla olan verilerde çok kısa ve çok uzun metinler veri setine katılmamıştır.

2.3. Veri Toplama Araçları

Karmaşıklık, doğruluk ve akıcılığın belirlenmesinde veri toplama aracı kullanılmamıştır. Bu konularda yapılan analizler sözcük, cümle sayılarına dayanmaktadır. Sayımlar birinci araştırmacı tarafından yapıldıktan sonra ikinci araştırmacı tarafından incelenmiştir. Ortak görüşte uzlaşma ile son karara varılmıştır.

İşlevsel yeterliliğin belirlenmesinde ise Kuiken ve Vedder (2017) tarafından güvenilirlik çalışmaları yapılan işlevsel yeterlik dereceli puanlama anahtarı kullanılmıştır. Dereceli puanlama anahtarı dört boyuttan oluşmaktadır. Bu boyutlar içerik, görev gereksinimleri, anlaşılabilirlik, tutarlık ve bağdaşıklık. Her boyut 1-6 puan arasında puanlanmaktadır. Ölçekten alınabilecek en yüksek puan 24’tür. Ölçeğin güvenilirlik çalışmaları L2 Hollandaca ve L2 İtalyanca öğrenen üniversite öğrencileri ile yapılmıştır. Hollandaca güvenilirlik puanları tüm boyutlarda .824-.940 ve İtalyanca için .725-.901 arasındadır. Türkçe için Özdemir (2024) tarafından yapılan güvenilirlik analizinde güvenilirlik puanlarının .838 ile .945 arasında olduğu belirlenmiştir. Bu çalışmada Türkçe öğretim merkezinde görev yapan ve öğrencilerin yazılı metinlerini puanlama deneyimi olan iki puanlayıcı tarafından puanlamalar yapılmıştır. Puanlayıcılara işlevsel yeterlik puanlama ölçütleri hakkında bilgi verilmiştir. Puanlayıcılar arası güvenilirlik Kappa analizi ile incelenmiştir. Puanlayıcılar arası uyum puanları içerik boyutunda .771, görev gereksinimleri .774, anlaşılabilirlik .736, tutarlık ve bağdaşıklık .846 olarak belirlenmiştir, puanlar arasındaki ilişki anlamlıdır ($p < .01$).

2.4. Veri Toplama Süreci

Araştırmanın verileri 2023-2024 eğitim ve öğretim yılında, bir kur süreci içinde toplanmıştır. Araştırmanın yürütüldüğü Bartın Üniversitede kurlar 6 hafta sürmektedir. Veriler birinci araştırmacı tarafından toplanmıştır. Araştırma için etik kurul onayı alınmıştır.

Öğrenciler, ders kitabı olarak kullanılan Yeni İstanbul Uluslararası Öğrenciler İçin Türkçe Ders Kitabı içinde yer alan yazma konularından ve yönergelerinden hareketle metinler yazmışlardır. Metinlerin konuları A2 düzeyinde sağlıklı yaşam, tatil planı ve ünlü birinin tanıtımıdır; B2 düzeyinde sosyal medyanın yararları ve zararları, mimari eser tanıtımı ve yenilenebilen enerji ile yenilenemeyen enerjinin karşılaştırmasıdır. Sonuç olarak her bir düzeyde 36 metin yazılmıştır, toplam 72 metin incelenmiştir.

ChatGPT'den her konuda birer metin yazması istenmiştir. Ancak A2 düzeyindeki sağlıklı yaşam konusunda öğrenciler "sigaranın zararları", sağlıklı beslenme" ve "uyku problemi yaşayanlar için tavsiyeler" şeklinde sunulan 3 seçenektan birini yazmışlardır. ChatGPT'den 3 alt konuda da metin hazırlaması istenmiştir. Sağlıklı yaşam konusundaki metinlerin ortalama puanları dikkate alınmıştır, diğer beş konuda ChatGPT sadece birer metin yazmıştır. Araştırmada ücretsiz erişilebilen ChatGPT 3.5 sürümü kullanılmıştır.

ChatGPT'ye verilen komutlar, oluşacak metnin niteliklerini etkileyebilir. Araştırmada ders kitabındaki yönergeler komut olarak kullanılmıştır. Sadece düzey konusunda ek bilgi verilmiştir, başka bir ek yönlendirme yapılmamıştır. Bu ilk komutla ortaya çıkan metinler incelenmiştir. Bu şekilde ortaya çıkacak metnin öğrenci metinleriyle daha uyumlu olabileceği göz önünde bulundurulmuştur. Örneğin tatil planları konusunda verilen yönerge şöyledir: Ben A2 düzeyinde Türkçe öğrenen yabancı uyruklu bir öğrenciyim. Tatil planı konusunda yazma çalışması yapmak istiyorum. Aşağıdaki soruları kullanarak benim için detaylı bir tatil planı yazar mısınız?

1. Nasıl bir tatil planlıyorsunuz? (Deniz turizmi/kültür turizmi/spor turizmi)
2. Tatile nereye gidiyorsunuz? (Yurt dışı/şehir dışı/şehre yakın bir yer)
3. Tatile ne zaman gideceksiniz? (Yazın, kışın)
4. Tatilde nerede kalacaksınız? (ev/otel/çadır)
5. Tatile nasıl gideceksiniz? (kendi aracım/kiralık araba/uçak/otostop)
6. Tatile kimlerle gideceksiniz? (yalnız/ailemle/arkadaşlarımla)
7. Tatilde kaç gün kalacaksınız? (hafta sonu/bir hafta-on gün/günübirlik)
8. Tatilde hangi aktiviteleri yapacaksınız? (tarihi yerleri gezmek, yüzmek, güneşlenmek, dalış yapmak, balık tutmak, arkadaşlarımla vakit geçirmek, bisiklete binmek, fotoğraf çekmek vb.)

2.5. Veri Çözümlemesi

Araştırmada karmaşıklık, doğruluk, akıcılık ve işlevsel yeterlik temalarında metinler analiz edilmiştir. Bulgular frekans ve yüzde çizelgeleri halinde sunulmuştur. Karmaşıklık, doğruluk, akıcılık açısından yapılan analizler için öncelikle toplam sözcük sayısı, hatalı ve hatasız sözcük sayıları, toplam cümle sayıları, hatalı ve hatasız cümle sayıları, basit ve karmaşık cümle sayıları belirlenmiştir. Öğrencilerin her konuda kendilerini tanıtarak giriş yaptıkları bölümler sözcük ve cümle sayısına dâhil edilmemiştir.

Öğrenci metinleri her konuda 12'şer tanedir. Öğrenci metinlerinin tamamından tek bir puan elde etmek için yazılan metinlerin toplamından elde edilen puan 12'ye bölünmüştür. Böylece düzeyin ortalama puanları belirlenmiştir. Hesaplamalarda bu ortalama puanlar kullanılmıştır. Öğrenci metinleri ve ChatGPT metinleri aşağıdaki ölçütlere göre ayrı ayrı incelenmiştir.

Karmaşıklık Analizi

Karmaşıklık cümle düzeyinde incelenmiştir. Basit cümle dışındaki bütün cümleler karmaşık cümle sayılmıştır. Karmaşıklığın belirlenmesinde basit cümleler ile basit cümle dışındaki cümle sayıları belirlenmiştir. Öğrencilerin bitmiş cümle sonunda virgül kullanmaları, bitmemiş cümle veya ifadeler sonunda nokta kullanmaları dikkate alınmamıştır. Cümle başında bağlaç varsa, kendinden önceki cümle ile anlamca bütünlük sağladığı için bu cümleler ayrı bir cümle olarak sayılmamıştır. Karmaşık cümle oranını belirlemek için “karmaşık cümle sayısı / toplam cümle sayısı x 100” formülü kullanılmıştır.

Doğruluk Analizi

Doğruluğu belirlemek üzere sözcük doğruluğu ve cümle doğruluğu incelenmiştir. Sözcük ve cümle hatalarını belirleme sürecinde dikkate alınan ölçütler aşağıda ilgili başlıkların altında sunulmuştur.

Sözcük doğruluğunun hesaplanması

Sözcük doğruluğunun hesaplanabilmesi için öncelikle hatalı sözcükler belirlenmiştir. Hatalı sözcükler belirlenirken sözcüğün anlamca uygun olmaması, yazımında harflerinin yanlış olması, yapım eklerinin yanlış olması ve başka dilden sözcük kullanılması ölçütleri dikkate alınmıştır. Büyük harf, küçük harf, şapkalı harf gibi yazım yanlışları ile ses bilgisine dayalı yanlışlar dikkate alınmamıştır.

Hatalı sözcüklerin sayısı belirlendikten sonra toplam sözcük sayısından bunlar düşülmüştür ve doğru sözcük sayısı elde edilmiştir. Bundan sonra doğru sözcük oranı hesaplanmıştır. Bunun için “doğru sözcük sayısı / toplam sözcük sayısı x 100” formülü kullanılmıştır.

Cümle doğruluğunun hesaplanması

Cümle doğruluğunu belirlemek için doğru ve hatalı cümle sayıları belirlenmiştir. Bu aşamada söz dizimi uygunsuzluğu, öge eksikliği, gereksiz sözcük bulunması, özne ile yüklem uyumsuzluğu, cümlelerin mantıksız olması, çekim eklerinin yanlış kullanılması veya kullanılması gerektiği halde ekin kullanılmaması gibi ölçütler dikkate alınarak yapıca ve anlamca hatalı cümlelerin sayısı belirlenmiştir.

Hatalı cümleler toplam cümle sayısından çıkarılarak doğru cümle sayısı belirlenmiştir. Cümle düzeyinde doğruluk oranını belirlemek için “doğru cümle / toplam cümle x 100” formülü kullanılmıştır.

Akıcılık Analizi

Akıcılığın belirlenmesi için öncelikle metindeki sözcük sayısı belirlenmiştir, ardından farklı sözcük sayısı ve sözcük yoğunluğu açısından incelemeler yapılmıştır. İncelemelerde dikkate alınan ölçütler aşağıda sunulmuştur.

Sözcük sayısını belirleme

Sözcük sayısı belirlenirken şu ölçütlere dikkat edilmiştir:

- Aynı kökten türeyen sözcükler ayrı bir sözcük olarak sayılmıştır.
- Fiilimsilerle ve olumsuzluk ekiyle türetilen sözcükler, soru eki -mI sözcük olarak sayılmıştır. Ergin (1985) soru anlamı veren –mI ekinin ad çekimi eklerinden sonra da kullanılabilmesi nedeniyle edat kökenli olduğunu belirtmektedir. Ediskun (2010) soru edatı kabul etmektedir. Soru eki, ayrı yazılması ve kendinden sonra ek alabilmesi nedeniyle ayrı bir sözcük olarak sayılmıştır.

• “ile” ve “ise” sözcükleri kendinden önceki sözcükle birleşik yazılsa bile ayrı bir sözcük olarak sayılmıştır.

• Ayrı yazılan birleşik fiiller (Ör. yardım+etmek, fotoğraf+çekmek), ayrı yazılan birleşik isimler (Ör. sosyal+medya, hafta+sonu), ayrı yazılan geçiş ve bağlantı ifadeleri (Ör. ya+da, son+olarak, bana+göre), özel isim ile birlikte anılan mekân adları (Ör. XX+Üniversitesi, Livingston+müzesi), ikilemeler (Ör. sık+sık), kişi ad ve soyadları (Ör. Kerem+Bürsin) aralarına artı işareti konarak tek sözcük sayılmıştır. Sözcüklerin yazımında Türk Dil Kurumu Güncel Türkçe Sözlük kaynak alınmıştır.

• Rakamla yazılan sayılar, kısaltmalar, yabancı ülke (Ör. Gabon), şehir (Ör. Java), mekân (Ör. Gedung sate), para birimi (Ör. rupiah) ve yemek (Ör. seblak) adları sözcük sayısına dâhil edilmemiştir.

• Deyimler kalıplaşmış sözler olsa da yabancı öğrencilerin birden çok sözcüğü yapıca ve anlamca uygun şekilde kullanması zor bir beceri olduğu için deyimler tek sözcük olarak sayılmamıştır.

Bu sayımlar sonucunda toplam sözcük sayısı belirlenmiştir.

Farklı sözcük sayısını belirleme

Farklı sözcük sayısı belirlenirken metinde tekrar eden sözcükler bir kez sayılmıştır. Örneğin sosyal medya konusunda yazılan bir metinde bu sözcükler 12 kez kullanılmışsa farklı sözcük sayısı olarak 1, aynı sözcük sayısı olarak 11 alınmıştır. Bu şekilde tekrar eden sözcükler toplanarak aynı sözcük sayısı belirlenmiştir. Aynı sözcükler toplam sözcük sayısından çıkarılarak farklı sözcük sayısı belirlenmiştir. Farklı sözcük sayısı belirlendikten sonra farklı sözcük oranı hesaplanmıştır. Farklı sözcük oranını hesaplamak için “farklı sözcük sayısı / toplam sözcük sayısı x 100” formülü kullanılmıştır.

Sözcük yoğunluğunu belirleme

Sözcük yoğunluğunu belirlemek için anlamlı ve görevli sözcüklerin sayısı belirlenmiştir. Görevli sözcükler konusunda Korkmaz (2009), tarafından yapılan sınıflama esas alınmıştır. Korkmaz’a göre isimler ve fiiller anlamlı sözcüklerdir. İsimler kapsamında isimler, sıfatlar, zamirler ve zarflar yer almaktadır. Görevli sözcükler edatlar ve bağlaçlardır. Ünlemler, kullanım durumuna göre görevli veya anlamlı sözcük özelliği taşıyabilmektedir. Ünlemler yalnız başına anlamlı ise, bir ad gibi çekimlenmişse veya türetilmişse anlamlı sözcük işlevi görmektedir. Çekime girmeyen, doğrudan bir kavrama karşılık gelmeyen ünlemler ise görevli sözcüklerdir. Bu çalışmada edatlar, bağlaçlar ve çekime girmeyen, türetme eki alarak isim özelliği taşımayan ünlemler görevli sözcükler olarak sayılmıştır.

Sözcük yoğunluğunu belirlemek için “anlamlı sözcük sayısı / toplam sözcük sayısı x 100” formülü kullanılmıştır.

Karmaşıklık, doğruluk ve akıcılık ölçütleri çerçevesinde incelemeler birinci araştırmacı tarafından yapıldıktan sonra ikinci araştırmacı tarafından kontrol edilmiştir. İki araştırmacının yaptıkları belirlemelerde farklılık olan durumlar üzerinde tartışılarak uzlaşma sağlanmıştır.

İşlevsel Yeterlik Analizi

İşlevsel yeterliliğin belirlenmesinde Kuiken ve Vedder (2017) tarafından geliştirilen işlevsel yeterlik dereceli puanlama anahtarı kullanılmıştır. Dereceli puanlama anahtarı içerik, görev gereksinimleri, anlaşılabilirlik, tutarlık ve bağdaşıklık boyutlarından oluşmaktadır. Her boyut 1-6 puan arasında puanlanmıştır. Puanlar 24’e yaklaştıkça işlevsel yeterliğin yüksek, 1’e yaklaştıkça düşük olduğu anlamındadır. Metinler yazma puanlaması yapmada uzman öğreticiler tarafından puanlanmıştır, puanlayıcıların verdiği puanların ortalaması alınmıştır. İşlevsel yeterlik puanı hesaplanırken öğrencilerin dört boyutta aldıkları puan toplamı işlevsel yeterlik puanını oluşturmuştur.

2.6. Etik İlkeler

Araştırmanın etik kurul izni Bartın Üniversitesi Sosyal ve Beşeri Bilimler Etik Kurulunun 16.11.2023 tarihli ve 19 numaralı kararı ile 2023-SBB-0747 numaralı başvuruya yönelik olarak alınmıştır. Katılımcılar araştırma hakkında bilgilendirilmiştir ve araştırmaya katılımın gönüllülüğe bağlı olduğu, araştırmaya katıldıktan sonra da istedikleri zaman araştırmadan çekilebilecekleri belirtilmiştir. Bu bilgilendirmenin yer aldığı onam formunu imzalayan katılımcılar ile araştırma yürütülmüştür.

3. Bulgular

Araştırmanın problem cümlesi öğrenciler ve ChatGPT tarafından üretilen metinlerin karmaşıklık, doğruluk, akıcılık ve işlevsel yeterlik açısından niteliklerinin belirlenmesine yöneliktir. Bu doğrultuda yapılan analizler aşağıda sunulmuştur.

Karmaşıklık Bulguları

Karmaşıklığın belirlenmesinde basit cümle dışındaki cümle yapıları karmaşık cümle sayılmıştır. Karmaşıklığa yönelik bulgular Çizelge 2’de sunulmuştur.

Çizelge 2. Öğrenciler ve ChatGPT metinlerinin karmaşıklık açısından durumu

Düzye	Konu	Öğrenci Ortalaması			ChatGPT				
		Basit	Karmaşık	Toplam	KCO*	Basit	Karmaşık	Toplam	KCO
A2	Sağlıklı Yaşam	3.3	5.9	9.2	64.1	4	15	19	78.9
	Tatil Planı	7	6.7	13.7	48.9	6	11	17	64.7
	Ünlü Tanıtımı	3.9	5.8	9.7	59.3	7	2	9	22.2
Ortalama KCO					57.4	55.3			
B2	Sosyal Medya	5	9.5	14.4	65.5	10	12	22	54.6
	Yer Tanıtımı	6.4	6.7	13.1	51.2	6	17	23	73.9
	Enerji	4.3	9.4	13.7	68.6	8	16	24	66.7
Ortalama KCO					61.8	65.1			

KCO*: Karmaşık cümle oranı

Çizelge 2’de görüldüğü gibi karmaşık cümle oranları A2 düzeyinde sağlıklı yaşam ve tatil planı konularında, B2 düzeyinde ise yer tanıtımı konusunda ChatGPT’nin ürettiği metinlerde öğrenci metinlerinden yüksektir. A2 düzeyinde ünlü tanıtımı konusunda öğrenci metinlerinde (%59.3) bu oran ChatGPT metninden (%22.2) daha yüksek iken B2 düzeyinde sosyal medya konusunda öğrenci metinlerinde (%65.5) ChatGPT metninden (%54.6) yüksek olduğu görülmektedir. B2 düzeyindeki enerji konusunda öğrenci metinleri ile ChatGPT metinlerinin karmaşıklık düzeyleri oldukça yakındır. Ortalama karmaşık cümle oranları A2 düzeyinde öğrenci metinlerinde %57.4’tür, ChatGPT metinlerinde %55.3’tür; B2 düzeyinde öğrenci metinlerinde %61.8’dir, ChatGPT metinlerinde %65.1’dir. ChatGPT’nin düzeylere göre tutarlı bir karmaşık cümle yapısını dikkate almadığı görülmektedir.

Doğruluk Bulguları

Doğruluğu belirlemek üzere cümle doğruluğu ve sözcük doğruluğu incelenmiştir. Cümle doğruluğu bulguları Çizelge 3’te sunulmuştur.

Çizelge 3. Öğrenciler ve ChatGPT metinlerinde cümle doğruluğu

Düzyey	Konu	Öğrenci Ortalaması				ChatGPT			
		Doğru	Hatalı	Toplam	DCO*	Doğru	Hatalı	Toplam	DCO
A2	Sağlıklı Yaşam	6.2	2.8	9.2	69.6	19	0	19	100
	Tatil Planı	4.6	9.1	13.7	33.6	17	0	17	100
	Ünlü Tanıtımı	4.1	5.6	9.7	42.3	9	0	9	100
Ortalama DCO					48.5				
B2	Sosyal Medya	2.6	11.9	14.5	18	22	0	22	100
	Yer Tanıtımı	4.8	8.3	13.1	36.7	21	1	23	95.7
	Enerji	5.38	7.9	13.7	42.3	24	0	24	100
Ortalama DCO					32.3	98.6			

DCO*: Doğru cümle oranı

Çizelge 3'te görüldüğü gibi öğrenciler tarafından üretilen metinlerin cümle düzeyinde doğruluk oranı %18 ile %69.6 arasında değişmektedir. ChatGPT tarafından üretilen metinlerde ise B2 düzeyinde yer tanıtımı konusundaki metinde doğruluk oranı %95.7, diğer metinlerin tamamında %100'dür. Ortalama doğru cümle oranları dikkate alındığında öğrencilerin ürettikleri cümlelerin yarıdan fazlasının hatalı olduğu dikkati çekmektedir. Öğrenci metinlerinde ortalama doğru cümle oranları dikkate alındığında A2 düzeyinde %48.5, B2 düzeyinde %32.3'tür. A2 düzeyinde doğru cümle oranının B2 düzeyinden daha fazla olması dikkat çekicidir.

ChatGPT'nin yer tanıtımı konulu metninde aşağıdaki cümlenin hatalı olduğu belirlenmiştir.

“İç mekânda sade ama şık bir tasarım görülür, detaylar genellikle geometrik desenler ve Arap harfleriyle bezeli olmuştur.”

Yukarıdaki cümlenin yüklemi uygun olmadığı için hatalı sayılmıştır. Cümlenin Türkçe ifadesi “(...) detaylar genellikle geometrik desenler ve Arap harfleriyle sağlanmıştır.” şeklinde olabilirdi.

Doğruluğun belirlenmesinde dikkate alınan diğer ölçüt sözcük doğruluğudur. Sözcük doğruluğu bulguları Çizelge 4'te sunulmuştur.

Çizelge 4. Öğrenciler ve ChatGPT metinlerinin sözcük düzeyinde doğruluğu

Düzyey	Konu	Öğrenci Ortalaması				ChatGPT			
		Doğru S.	Hatalı S.	Top.	DSO*	Doğru S.	Hatalı S.	Top. S.	DSO
A2	Sağlıklı Yaşam	92.7	2.6	95.3	97.3	61	0	61	100
	Tatil Planı	75.7	10.4	86.1	88	109	0	109	100
	Ünlü Tanıtımı	66.3	5.2	71.5	92.7	50	0	50	100
Ortalama DSO					92.7	100			
B2	Sosyal Medya	128.1	13.9	142	90.2	238	0	238	100
	Yer Tanıtımı	97.6	8.7	106.3	91.8	272	6	278	97.8
	Enerji	123	10.1	133.1	92.4	284	0	284	100
Ortalama DSO					91.5	99.3			

DSO*: Doğru sözcük oranı

Çizelge 4'te sözcük düzeyinde doğruluk açısından öğrenci metinlerinde doğruluk oranının %88 ile %97.3 arasında değiştiği görülmektedir. A2 düzeyinde ortalama doğruluk oranı %92.7'dir, B2 düzeyinde %91.5'tir. Ortalama doğru sözcük oranının A2 düzeyinde B2 düzeyinden yüksek olduğu görülmektedir. ChatGPT metinlerinde ise sözcüksel doğruluk B2 düzeyindeki yer tanıtımı konulu metinde %97.8'dir, diğer bütün metinlerde %100'dür. Ortalama doğruluk oranı A2 düzeyinde %100 iken B2 düzeyinde %99.3'tür. ChatGPT'nin az sayıda da olsa anlamca uygun olmayan sözcükler kullanması dikkat çekicidir. ChatGPT

tarafından oluşturulan metinde sözcüklerin anlamca uygun olmadığını örnekleyen bir cümle aşağıda sunulmuştur:

“*Bartın+Ulu+Cami'nin iç kısmı, oldukça geniş bir alan sunar.*” cümlesindeki “sunar” yüklemi yerine “(...) oldukça geniş bir alana sahiptir.” ifadesi Türkçe için daha uygundur.

Akıcılık Bulguları

Akıcılığı belirlemek üzere tekrarlardan arınmış farklı sözcük kullanımı ve sözcük yoğunluğu açısından karşılaştırmalar yapılmıştır. Farklı sözcük kullanımına yönelik bulgular Çizelge 5’te sunulmuştur.

Çizelge 5. Öğrenciler ve ChatGPT metinlerinde farklı sözcük kullanımı

Düzye	Konu	Öğrenci Ortalaması			ChatGPT				
		Aynı S.	Farklı S.	Top.	FSO*	Aynı S.	Farklı S.	Top.	FSO
A2	Sağlıklı Yaş.	22.4	72.9	95.3	76.5	10.7	50.3	61	82.5
	Tatil Planı	24.3	61.8	86.1	71.8	33	76	109	69.7
	Ünlü Tanıt.	13.5	58	71.5	81.1	6	44	50	88
Ortalama FSO					76.5	80.1			
B2	Sosyal Medya	34	92	142	64.8	59	179	238	75.2
	Yer Tanıt.	35.6	70.7	106.3	66.5	97	181	278	65.1
	Enerji	44.8	88.3	133.1	66.3	99	185	284	65.1
Ortalama FSO					65.9	68.5			

FSO*: Farklı sözcük oranı

Çizelge 5 incelendiğinde farklı sözcük kullanımının A2 düzeyinde (%76.5) B2 düzeyinden (%65.9) yüksek olduğu görülmektedir. Öğrenciler ile ChatGPT metinleri karşılaştırıldığında ChatGPT metinlerinin her iki düzeyde de farklı sözcük oranlarının daha yüksek olduğu görülmektedir. ChatGPT metinlerinde farklı sözcük kullanma oranı A2 düzeyinde %80.1 ve B2 düzeyinde %68.5’tir.

Akıcılığın belirlenmesinde dikkate alınan ölçütlerden biri de sözcük yoğunluğudur. Sözcük yoğunluğu anlamlı ve görevli sözcük kullanımına dayalı olarak belirlenmiştir. Sözcük yoğunluğuna yönelik bulgular Çizelge 6’da sunulmuştur.

Çizelge 6. Öğrenciler ve ChatGPT metinlerinde sözcük yoğunluğu

Düzye	Konu	Öğrenci Ortalaması				ChatGPT			
		Görevli S.	Anlamlı S.	Top.	ASO*	Görevli S.	Anlamlı S.	Top.	ASO
A2	Sağlıklı Yaşam	10.7	84.6	95.3	88.8	7.3	53.7	61	88
	Tatil Planı	9.4	76.6	86.1	90.1	15	94	109	86.2
	Ünlü Tanıtımı	8.2	63.3	71.5	89	5	45	50	90
Ortalama					89.3	88.1			
B2	Sosyal Medya	23.1	118.9	142	83.7	43	195	238	81.9
	Yer Tanıtımı	12.3	94	106.3	88.4	33	245	278	88.1
	Enerji	19.8	113.3	133.1	85.1	32	252	284	88.7
Ortalama					85.7	85.2			

ASO*: Anlamlı sözcük oranı

Çizelge 6’da görüldüğü gibi anlamlı sözcük oranı ortalamaları A2 düzeyindeki metinlerde (%89.3) B2 düzeyindeki metinlerden (%85.7) daha yüksektir. Benzer şekilde ChatGPT’nin ürettiği metinlerde de A2 düzeyindeki metinlerde (%88.1) B2 düzeyindeki metinlerden (%85.2) daha fazla anlamlı sözcük

kullanıldığı görülmektedir. Öğrenciler ile ChatGPT'nin ürettiği metinler karşılaştırıldığında her iki düzeyde de öğrenci metinleri nispeten daha fazla anlamlı sözcük içermektedir.

İşlevsel Yeterlik Bulguları

Araştırmada metinlerin niteliklerinin karşılaştırılmasında kullanılan bir ölçüt de işlevsel yeterliktir. Metinlerin işlevsel yeterliğini belirlemek amacıyla yapılan analizler Çizelge 7'de sunulmuştur.

Çizelge 7. Öğrenciler ve ChatGPT metinlerinde işlevsel yeterlik

Düzye	Konu	Öğrenci Ortalaması					ChatGPT				
		İçerik	Görev Gereksinimleri	Anlaşılabilirlik	Tutarlılık ve Bağdaşlılık	İşlevsel Yeterlik Puanı	İçerik	Görev Gereksinimleri	Anlaşılabilirlik	Tutarlılık ve Bağdaşlılık	İşlevsel Yeterlik Puanı
A2	Sağlıklı Yaşam	4.6	4.8	4.9	4.6	18.9	6	6	6	6	24
	Tatil Planı	4.5	4.9	4.5	4.3	18.2	6	6	6	6	24
	Ünlü Tanıtımı	3.9	4.4	3.9	4	16.2	4	4.5	5.5	5	19
Ortalama İşlevsel yeterlik Puanı					17.8						22.3
B2	Sosyal Medya	4.6	4.5	4.8	4.6	18.8	6	6	6	6	24
	Yer Tanıtımı	3.5	3.4	3.9	3.5	14.3	6	6	6	6	24
	Enerji	3.8	3.3	3.8	3.5	14.4	6	6	6	6	24
Ortalama İşlevsel yeterlik Puanı					15.8						24

Çizelge 7'de görüldüğü gibi işlevsel yeterlik ortalama puanları A2 düzeyinde 17.8 iken B2 düzeyinde 15.8'dir. Temel düzeydeki öğrencilerin orta düzeydeki öğrencilerden daha nitelikli metinler yazdıkları görülmektedir. ChatGPT metinleri iki düzeyde de öğrenci metinlerinden daha yüksek işlevsel yeterlik puanına sahiptir. ChatGPT metinlerinin A2 düzeyinde işlevsel yeterlik ortalama puanı 22.3'tür, B2 düzeyinde tüm ölçütlerden tam puan olarak toplam 24 puana eriştiği görülmektedir. ChatGPT'nin ürettiği A2 düzeyindeki ünlü tanıtımı konulu metin 50 sözcük ve 9 cümleden oluşmaktadır. Az sayıda sözcük ve cümle ile oluşturulan metnin işlevsel yeterliğinin de diğer metinlerden az puana sahip olduğu dikkati çekmektedir.

4. Tartışma, Sonuç ve Öneriler

Bu çalışmada, Türkçeyi yabancı dil olarak öğrenen öğrencilerin yazdıkları metinler ile ChatGPT tarafından oluşturulan metinler karmaşıklık, doğruluk, akıcılık ve işlevsel yeterlik ölçütleri doğrultusunda karşılaştırılmıştır. Araştırmanın başlıca sonuçları ve bu sonuçlara yönelik tartışmalar aşağıda sunulmuştur.

Karmaşıklık açısından değerlendirildiğinde, genel olarak, ChatGPT metinlerinin öğrenci metinlerine göre daha yüksek karmaşık cümle oranına sahip olduğu görülmüştür. Ancak A2 seviyesinde ünlü tanıtımı ve B2 seviyesinde sosyal medya konularında öğrenci metinlerinin karmaşık cümle oranının ChatGPT metinlerinden daha yüksek olduğu tespit edilmiştir. A2 seviyesindeki ortalama karmaşık cümle oranı öğrenci metinlerinde %57.4 iken, ChatGPT metinlerinde %55.3'tür. B2 seviyesinde ise bu oran öğrenci metinlerinde %61.8 iken, ChatGPT metinlerinde %65.1'dir.

A2 düzeyinde ortalama karmaşık cümle oranının öğrenci metinlerinde ChatGPT metinlerinden yüksek olması, temel düzeydeki öğrencilerin beklenenden daha karmaşık yapılar kullanma eğiliminde olduklarını düşündürmektedir. Bu çalışma kapsamında, öğrencilerin dil bilgisi yeterliliğinin yanı sıra, metin

türü ve konusunun karmaşıklık düzeyini etkilediği gözlenmiştir. Örneğin, “ünlü tanıtımı” gibi bir konuda, “...olan bir kişi”, “...yapan bir yazar” gibi ifadelerin veya “...dan dolayı...” gibi bağlaçların sıkça kullanılması karmaşık yapıları doğal olarak oluşabilmektedir. B2 düzeyinde ChatGPT metinlerinde karmaşık cümle oranı öğrenici metinlerinden yüksek çıkmıştır. ChatGPT’nin öğrenici düzeyine uygun metin üretme konusunda tutarlı bir yaklaşımının olmadığı söylenebilir. Zileli (2023) tarafından yapılan araştırmada da ChatGPT’nin Türkçe dil seviyesine uygun cümle yapıları kullanmadığı belirlenmiştir. Öğrenici ve ChatGPT metinlerini karşılaştıran araştırmalarda İngilizcede de (Freerick ve Craven, 2025; Mizumoto vd., 2024) ChatGPT’nin karmaşık cümle üretmede daha başarılı olduğu belirlenmiştir. Araştırmaların ortak sonuçları değerlendirildiğinde ChatGPT’nin dil düzeyleri açısından geliştirilmeye ihtiyacı olduğu söylenebilir.

Doğruluk boyutunda elde edilen bulgular, ChatGPT’nin hem cümle hem de sözcük düzeyinde oldukça yüksek doğruluk oranlarına ulaştığını göstermektedir. Cümle doğruluğu açısından, B2 seviyesindeki yer tanıtımı metni (%95.7) dışında tüm ChatGPT metinlerinde doğruluk oranı %100 iken, öğrenici metinlerinde bu oran %18 ile %69.6 arasında değişmektedir. Öğrenici metinlerinde ortalama doğru cümle oranı A2 düzeyinde %48.5 iken B2 düzeyinde %32.3’tür.

Öğrencilerin doğru cümle yazmada büyük sorunlar yaşadıkları görülmektedir. İncelenen metinlerde cümlelerin yarıdan fazlası hatalıdır. ChatGPT’nin ise hem cümle hem de sözcük düzeyinde neredeyse kusursuz bir doğruluk performansı sergilediği belirlenmiştir. Yabancı dil öğreniminde, öğrencilerin yazılı anlatımlarında sıklıkla dil bilgisel hatalar, sözcük seçimi sorunları ve bağlaç uyumsuzlukları gibi problemlerle karşılaştıkları gözlemlenmektedir (Demirel ve Dağlı, 2016; Özdemir ve Şimşek, 2025). Ancak öğrenici metinlerindeki cümlelerin yarıdan fazlasının hatalı olması, hataların düzey ilerledikçe artması, A2 düzeyinde %48.5, B2 düzeyinde %32.3 oranında doğru cümle bulunması kabul edilebilecek bir durum değildir. Bu durum, uygulanan programın içeriğinin öğrenciler tarafından edinilemediğini düşündürmektedir. Öğrencilerin ek materyallerle desteklenmesi yararlı olabilir. Dil bilgisel doğruluğun geliştirilmesi uygulamalarında ChatGPT de güçlü bir destek sunabilir. ChatGPT’nin dil bilgisel hataları otomatik olarak düzeltme yeteneği, düzeltici geri bildirim mekanizmalarına benzer bir işlev görebilir (Baleghizadeh ve Gordani, 2012; Yan, 2023).

Sözcüksel doğruluk oranları öğrenici metinlerinde ortalama %91.5, ChatGPT metinlerinde %99.3 olarak belirlenmiştir. ChatGPT’nin sözcük düzeyinde her zaman bağlama uygun sözcük seçemediği görülmüştür. Zileli (2023) tarafından yapılan araştırmada da ChatGPT’nin doğruluk konusunda bazı sınırlılıkları olduğu belirlenmiştir. Özellikle, metindeki anlamsal hataları tespit etmede yetersiz kalabildiği görülmüştür. Bu sonuçlar ChatGPT’nin sözcüksel doğruluk konusunda geliştirilmesi gerektiğini, dolayısıyla öğrencilerin bu durumlara karşı dikkatli olmaları gerektiğini göstermektedir. Bununla birlikte öğrenici metinlerinden daha iyi bir performans sergilediği için sözcüksel doğruluğun geliştirilmesinde yararlanılabilecek bir araç durumundadır. ChatGPT’nin ürettiği metinlerdeki bağlama uymayan sözcükler öğretim aracı haline getirilip tartışılırsa yeni sözcükler öğrenme konusunda gelişim sağlanabilir.

Akıcılığın değerlendirilmesinde farklı sözcük oranı ve anlamlı sözcük oranı açılarından incelemeler yapılmıştır. Farklı sözcük oranı açısından, her iki seviyede de ChatGPT metinlerinin öğrenici metinlerinden daha yüksek orana sahip olduğu görülmüştür. Farklı sözcük oranı öğrencilerde A2 düzeyinde %76.5, B2 düzeyinde %65.9 olarak belirlenmiştir. ChatGPT metinlerinde farklı sözcük oranı A2 düzeyinde %80.1 ve B2 düzeyinde %68.5 olarak belirlenmiştir. Yıldız Durak vd. (2025) ana dili Türkçe olan üniversite öğrencilerinin ürettikleri metinlerle ChatGPT metinlerini karşılaştırdıkları araştırmada farklı sözcük kullanımının öğrenci metinlerinde fazla olduğunu belirlemişlerdir. Bu durumda Türkçeyi yabancı dil olarak öğrenen öğrencilerin sözcük bilgisi daha sınırlı olduğu için ChatGPT’nin üstün olduğu söylenebilir. Yapay

zekâ tarafından üretilen metinlerin öğrenici metinlerinden daha yüksek sözcük çeşitliliğine ve akıcılığa sahip olduğu farklı araştırmalarda da görülmektedir (Fredrick ve Craven, 2025; Yan, 2023). Bu sonuçlar değerlendirildiğinde Türkçeyi yabancı dil olarak öğrenenlerin sözcük bilgisini geliştirmelerinde ChatGPT'nin destek sunabilecek nitelikte olduğu söylenebilir.

Anlamli sözcük oranı karşılaştırmasında ise, öğrenici metninin her iki seviyede de ChatGPT metnlerine çok yakın anlamli sözcük içerdiği görülmüştür. A2 seviyesindeki öğrenici metnlerinde anlamli sözcük oranı % 89.3 iken, ChatGPT metnlerinde bu oran %88.1'dir. B2 seviyesinde ise öğrenici metnlerinde anlamli sözcük oranı %85.7 iken ChatGPT metnlerinde %85.2'dir. Anlamli sözcükler kullanmak fiillerin ve isimlerin yerli yerinde kullanıldığını göstermektedir. Bütün dillerde olduğu gibi Türkçede de anlamli sözcüklerin sayısı görevli sözcüklerden fazladır (Sarı, 2000). Bu nedenlerle anlamli sözcük kullanma oranlarının yüksek olması doğaldır. Ancak bu çalışmada bazı öğrenici metninin anlamli sözcük oranı bakımından ChatGPT'den güçlü olduğu görülmüştür. Görevli sözcükler, anlamli sözcüklerin cümle oluşturmasını sağlayan, karmaşık cümle oluşturmaya destek sunan öğelerdir. Araştırmada özellikle B2 düzeyindeki öğrenici metninin karmaşıklık bakımından ChatGPT'nin gerisinde kaldıkları, her iki düzeyde öğrencilerin doğru cümle kurmakta çok zorlandıkları görülmüştür. Dolayısıyla öğrencilerin ChatGPT'den daha fazla anlamli sözcük kullanmaları, görevli sözcükleri kullanmaktan kaçınmaları, hatalı cümle kurmaktan korkmalarıyla ilişkili olabilir.

İşlevsel yeterlik açısından ChatGPT metninin öğrenici metinlerinden daha yüksek puanlara sahip olduğu belirlenmiştir. A2 seviyesinde ortalama işlevsel yeterlik puanı öğrencilerde %17.8 iken ChatGPT'de %22.3'tür. B2 seviyesinde ise öğrencilerde %15.8 iken ChatGPT'de %24'tür. Bu sonuçlar ChatGPT'nin sadece doğru cümle kurma becerisine değil, aynı zamanda tutarlı içerik üretme, görev gereksinimlerini karşılama ve metin bütünlüğü sağlama yeteneğine de sahip olduğunu göstermektedir. ChatGPT'nin konuya uygun, belli bir amaç etrafında tutarlı metinler üretebildiği Katı ve Can (2024) tarafından yapılan araştırmada da görülmüştür. Ancak bu araştırmada bazı durumlarda içerik derinliğinin eksikliği dikkat çekmektedir. Zhou vd. (2023) tarafından yapılan araştırmada tutarlılık açısından insan yazılarının daha üstün olduğu belirlenmiştir. Bu durum, yapay zekâ araçlarının psikolojik derinlikten uzak olduğunu göstermektedir. Bununla birlikte ChatGPT dil bilgisi ve görevin gerektirdiği içeriği tutarlı bir şekilde üretmede öğrencilerden daha başarılı olmuştur. Bu sonuçlar, ChatGPT'nin yazma eğitimini destekleyici bir biçimde, dengeli ve bilinçli bir şekilde kullanılması gerektiğini vurgulayan araştırmaları desteklemektedir (Mondal ve Mondal, 2023; Sallam, 2023).

Doğruluk, akıcılık ve işlevsel yeterlik puanlarının A2 düzeyinde B2 düzeyinden daha yüksek olması araştırmanın dikkat çeken bulgularından biridir. Öğrencilerin dil yeterlikleri geliştikçe doğru, akıcı ve işlevsel açıdan yeterli metinler yazması beklenir. Bu araştırmada ise A2 düzeyindeki öğrencilerin B2 düzeyindeki öğrencilerden daha iyi performans sergiledikleri görülmüştür. Bu durum program, materyaller, ders saatleri, öğretici yeterlikleri, öğrenici nitelikleri gibi çeşitli değişkenlerle ilişkili olabilir, ancak çalışma kapsamında bu değişkenler doğrudan test edilmemiştir. Bununla birlikte A2 grubundaki 21 öğrenciden 12'sinin Kazakistanlı olması ile ilgili bir durum olabileceği de düşünülmektedir. Kültürel ve dilsel yakınlık avantajına sahip olan bu öğrenciler nispeten homojen bir ortamda birbirleriyle daha kolay iletişim kurarak ve öğrenme stratejileri geliştirerek başarılarını artırmış olabilirler. Bu durum, gelecekteki çalışmalar için sınıf bileşiminin de önemli bir değişken olduğuna işaret etmektedir.

Sonuç olarak, bu araştırma ChatGPT'nin yabancılara Türkçe öğretiminde yazma becerisini geliştirme sürecinde cümle ve sözcük düzeyinde doğruluğu geliştirme ve anlam oluşturma açısından metnin işlevsel yeterliğinin sağlanmasında potansiyel bir destek aracı olabileceğini göstermektedir. İşlevsel yeterlik; metnin anlamı açısından yeterli ve konuyla ilgili bir içerik sunmasını, görev gereksinimlerini karşılmasını, amacı

ve fikirleri açıkça anlaşılan metin yazmayı, metni geçiş ve bağlantı ifadeleriyle tutarlı bir bütün haline getirmeyi kapsamaktadır. Yabancı dil olarak Türkçe öğrenenlerin dil bilgisi açısından doğru, anlam açısından türe uygun bir içerik düzenlemelerini desteklemek için ChatGPT'den yararlanılabileceği görülmektedir. Öğrencilerin özgün tasarımlarını yapmaları, yaratıcılıklarını korumaları, dilsel gelişimlerini sürdürmelerini için ChatGPT'den metni planlama veya yazılan metni gözden geçirme aşamalarında yararlanmak daha uygun olacaktır.

Bu araştırmanın bazı sınırlılıkları vardır. Bunlardan biri incelenen metinlerin betimleme, bilgilendirme, karşılaştırma gibi farklı işlevlere yönelik olmasıdır. Belli bir işleve yönelik öğrenci metinleri ile yapay zekâ metinlerinin karşılaştırılması daha derin bir anlayış elde etmemizi sağlayabilir. Araştırmanın bir diğer sınırlılığı karşılaştırmalı betimleme tasarımıyla yürütülmesidir. Bu tasarım metinler hakkında fikir vermektedir. Öğrencilerin yapay zekâdan yararlanarak yazma becerisi geliştirme konusunda pedagojik olarak nasıl etkilendikleri ise ayrıca araştırılmalıdır.

Araştırma kapsamında öğretim sürecine ve sonraki araştırmalara yönelik olarak aşağıdaki öneriler geliştirilmiştir.

Yabancılarla Türkçe öğretiminde yazma becerisini geliştirme sürecinde ChatGPT'den cümle ve sözcük doğruluğunu geliştirme ve metnin işlevsel yeterliğini sağlama açısından yararlanılabilir.

Öğrencilerin bireysel gelişimini sürdürebilmeleri amacıyla ChatGPT gibi araçlar özellikle yazma öncesi fikir üretme ve yazma sonrası düzeltme süreçlerinde kullanılabilir.

ChatGPT'nin sözcük seçiminde ve işlevsel yeterlikte her zaman mükemmel olmayabileceği dikkate alınmalı, ChatGPT metinlerinin eleştirel bir gözle değerlendirilmesi teşvik edilmeli ve öğrencilerin kendi içeriklerini üretmeleri sağlanmalıdır.

Metin türü ve dil düzeyi sınırlandırılarak benzer araştırmalar ChatGPT ve diğer üretken yapay zekâ araçları için planlanabilir.

Uygulama temelli araştırmalarla ChatGPT'nin pedagojik etkileri öğrenci ve öğretmen görüşlerine göre incelenebilir.

Etik Bildirimler

Bu dergi COPE (Committee on Publication Ethics) ilkelerini benimser.

Çıkar çatışması

Yazar(lar) herhangi bir çıkar çatışması bulunmadığını bildirir.

Finansman

Bu çalışma herhangi bir kurum/kuruluş tarafından desteklenmemiştir.

Kaynakça doğrulama bildirimini

Yazar(lar) kaynakçada yer alan tüm çalışmaların metin içinde atıf aldığını ve doğruluğunun denetlendiğini bildirir.

İntihal bildirimini

Bu çalışma intihal saptama yazılımları ile taranmış olup bilimsel etik kurallarına uygun hazırlanmıştır.

Hakemlik süreci

Bu makale çift kör hakemlik sürecinden geçirilmiştir.

Makale erişim ve lisans

Yazarlar çalışmalarının telif hakkını saklı tutar ve dergiye ilk yayımlama hakkını verir. Tüm makaleler Creative Commons Atıf 4.0 Uluslararası (CC BY 4.0) lisansı kapsamında lisanslanmaktadır.

Bilgilendirme

Bu çalışma birinci yazarın ikinci yazar danışmanlığında hazırladığı “Yabancı dil olarak Türkçe öğrencilerinin yazılı metinleri ile ChatGPT tarafından yazılan metinlerin niteliklerinin karşılaştırılması” adlı yüksek lisans tezinden üretilmiştir.

Kaynakça

- An, N. M. (2024). Leveraging ChatGPT for enhancing English writing skills and critical thinking in university freshmen. *Journal of Knowledge Learning and Science Technology*, 3(2), 51-62. <https://jklst.org/index.php/home/article/view/168/142>
- Baleghizadeh, S., & Gordani, Y. (2012). Academic writing and grammatical accuracy: The role of corrective feedback. *Gist Education and Learning Research Journal*, 6, 159-176.
- Balyemez, S. (2021). *Dil bilgisi üzerine açıklamalar*. (6. Baskı). Pegem Akademi.
- Baş, B. ve Turhan, O. (2017). Yabancılara Türkçe öğretiminde yazma becerisine yönelik Web 2.0 araçları: Poll Everywhere örneği. *Mersin Üniversitesi Eğitim Fakültesi Dergisi*, 13(3), 1233-1248.
- Bellman, R. (1978). *An introduction to artificial intelligence: Can computers think?*. Boyd & Fraser Pub.
- Bülbül, A., ve Karakaş, B. (2024). Uluslararası öğrencilerin yazılı metinlerinde tutarlılık görünümüleri. *Dil Dergisi*, 175(2), 116-148. <https://doi.org/10.33690/dilder.1525419>
- Cantrell, M. A. (2011). Demystifying the research process: Understanding a descriptive comparative research design. *Pediatric Nursing*, 37(4), 188-189.
- Celce-Murcia, M. (1991). Grammar pedagogy in second and foreign language teaching. *TESOL quarterly*, 459-480. <https://doi.org/10.2307/3586980>
- Council of Europe. (2020), *Common European framework of reference for languages: Learning, teaching, assessment- Companion volume*. Council of Europe Publishing. www.coe.int/lang-cefr.
- Dizdaroğlu, H. (1976). *Tümcebilgisi*. Türk Dil Kurumu Yayınları.
- Ediskun, H. (2010). *Türk dilbilgisi*. Remzi Kitabevi.
- Ergin, M. (1985). *Türk dil bilgisi*. Bayrak Yayıncılık.
- Fredrick, D. R., & Craven, L. (2025). Lexical diversity, syntactic complexity, and readability: a corpus-based analysis of ChatGPT and L2 student essays. *Frontiers in Education*, <https://doi.org/10.3389/feduc.2025.1616935>
- Gencan, T. N. (1979). *Dilbilgisi*. Türk Dil Kurumu Yayınları.
- Heller, C. H. (2019). Near-term applications of artificial intelligence: Implementation opportunities from modern business practices. *Naval War College Review*, 72(4), 73-100. <https://www.jstor.org/stable/26775520>
- Herbold, S., Hautli-Janisz, A., Heuer, U., Kikteva, Z., & Trautsch, A. (2023). A large-scale comparison of human-written versus ChatGPT-generated essays. *Scientific Reports*, 13:18617. <https://doi.org/10.1038/s41598-023-45644-9>
- Karahan, L. (2000). Yapı bakımından cümle sınıflandırmaları üzerine. *Türk Dili*, (583), 16-23.

- Katı, T. N. ve Can, U. (2024). Yapay zekâ ile üretilen metinlerin yabancı dil olarak Türkçe öğretiminde okuma becerisine yönelik kullanılabilirliği: ChatGPT-3.5 örneği. *İnönü Üniversitesi Eğitim Fakültesi Dergisi*, 25(2), 538-569. <https://doi.org/10.17679/inuefd.1415303>
- Korkmaz, Z. (2009). *Gramer terimleri sözlüğü*. Türk Dil Kurumu Yayınları.
- Koşaner, Ö., Çimen, B. ve Karacan, P. (2023). Dil modeli üretimi ve insan yazımı metinlerin karşılaştırmalı çözümlemesi: Akademik metinler bağlamında bir inceleme. İçinde V. Uğur (Ed.), *Günce filoloji ve sosyal bilimler çalışmaları I* (1. baskı, ss. 144-162). Günce Yayınları.
- Krippendorff, K. (2018). *Content analysis: An introduction to its methodology* (4th ed.). Sage.
- Kuiken, F. (2023). Linguistic complexity in second language acquisition. *Linguistics Vanguard*, 9(1): 83–93. <https://doi.org/10.1515/lingvan-2021-0112>
- Kuiken, F., & Vedder, I. (2017). Functional adequacy in L2 writing: Towards a new rating scale. *Language Testing*, 34(3), 321–336. <https://doi.org/10.1177/0265532216663991>
- Larsen-Freeman, D. (1978). An ESL index of development. *TESOL Quarterly*, 12(4), 439-448. <https://about.jstor.org/terms>
- Larsen-Freeman, D. (2006). The emergence of complexity, fluency, and accuracy. *Applied Linguistics*, 27(4), 590–619.
- Mizumoto, A., Yasuda, S., & Tamura, Y. (2024). Identifying ChatGPT-generated texts in EFL students' writing: Through comparative analysis of linguistic fingerprints. *Applied Corpus Linguistics*, 4(1), 100106. <https://doi.org/10.1016/j.acorp.2024.100106>
- Mondal, H., & Mondal, S. (2023). ChatGPT in academic writing: Maximizing benefits and minimizing risks. *Indian Journal of Ophthalmology*, 71(12), 3600-3606. https://doi.org/10.4103/IJO.IJO_718_23
- Özdemir, S. (2024). The effects of collaborative writing on L2 individual writing performance in terms of complexity, accuracy, fluency and functional adequacy. *Pegem Journal of Education and Instruction*, 14(4), 235-244. <https://doi.org/10.47750/pegegog.14.04.20>
- Özdemir, S. ve Şimşek, P. H. (2025, 20-22 Haziran). Yabancı dil olarak Türkçe öğrenen öğrencilerin yazılı anlatım hataları. *7. Başkent International Conference on Multidisciplinary Studies*, Ankara.
- Öztürk Demir, S. ve Çetin, M. (2025). Bir üretken yapay zekâ aracı olarak ChatGPT'nin yabancı dil öğretiminde kullanılmasına ilişkin görüşler. *Dil Eğitimi ve Araştırmaları Dergisi*, 11(1), 207-234. <https://doi.org/10.31464/jlere.1501596>
- Richards, J. C., & Renandya, W. A. (2002). *Methodology in language teaching: An anthology of current practice*. Cambridge University Press.
- Sallam, M. (2023). ChatGPT utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns. *Healthcare*, 11(6), 887. <https://doi.org/10.3390/healthcare11060887>
- Sarı, M. (2000). Türkiye Türkçesinde sözcük türlerinin değişikliğe uğraması. *Türkoloji Dergisi*, 13(1), 229-238. https://doi.org/10.1501/Trkol_0000000121
- Song, C., & Song, Y. (2023). Enhancing academic writing skills and motivation: assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. *Frontiers in Psychology*, 14:1260843. <https://doi.org/10.3389/fpsyg.2023.1260843>
- Storch, N. (2005). Collaborative writing: Product, process, and students' reflections. *Journal of Second Language Writing*, 14(2005), 153-173. <https://doi.org/10.1016/j.jslw.2005.05.002>
- Yan, D. (2023). The effect of ChatGPT on students' second language writing practice: An exploratory inquiry. *Education and Information Technologies*, 28, 13942-13967. <https://doi.org/10.1007/s10639-023-11742-4>

- Yıldız Durak, H., Eğin, F., & Onan, A. (2025). A comparison of human-written versus AI-generated text in discussions at educational settings: Investigating features for ChatGPT, Gemini and BingAI. *European Journal of Education*, 60:e70014. <https://doi.org/10.1111/ejed.70014>
- Zhai, X. (2022). ChatGPT user experience: Implications for education. <https://doi.org/10.2139/ssrn.4312418>
- Zhang, P., & Tur, G. (2024). A systematic review of ChatGPT use in K-12 education. *European Journal of Education*, 59(2), 1-22. <https://doi.org/10.1111/ejed.12599>
- Zhou, T., Cao, S., Zhou, S., Zhang, Y., & He, A. (2023). Chinese intermediate English learners outdid ChatGPT in deep cohesion: Evidence from English narrative writing. arXiv. <https://doi.org/10.48550/arXiv.2303.11812>
- Zileli, E. N. (2023). Yabancı dil olarak Türkçe öğreniminde ChatGPT örneği. *Karamanoğlu Mehmet Bey Uluslararası Eğitim Araştırmaları Dergisi*, 5(1), 42-51. <https://doi.org/10.47770/ukmead.1296013>

Extended Abstract

Complexity, Accuracy, Fluency, and Functional Adequacy in Texts Written by Learners of Turkish as a Foreign Language and Texts Generated by ChatGPT

Writing in a foreign language is a language skill that tangibly demonstrates linguistic development. However, linguistic competence supports the construction of grammatically correct and meaningful sentences, but does not guarantee the production of coherent and genre-appropriate texts. The development of writing skills is directly related to individual awareness and effective feedback; however, providing immediate feedback on writing errors is not always possible in traditional teaching environments. In this case, generative artificial intelligence tools have attracted attention with their human-like text generation capabilities and have become a potential auxiliary tool in language learning and teaching processes. However, artificial intelligence tools may not always be reliable because the results they produce depend on the data set they are trained on. Therefore, educators need to examine the adequacy of artificial intelligence in their own languages and subject areas before using it in their lessons.

The purpose of this study is to compare the qualities of texts written by learners of Turkish as a foreign language with those generated by ChatGPT. The research question is defined as follows: How do texts written by learners of Turkish as a foreign language and those generated by ChatGPT compare in terms of complexity, accuracy, fluency, and functional adequacy?

The research was conducted using a comparative descriptive research design. Forty-three learners at the A2 and B2 levels at Bartın University's Turkish Language Teaching Center participated in the study. The students wrote texts based on the writing topics and guidelines in the textbook. Twelve learners at each level wrote texts on each topic. A total of 36 texts were written at each level, and 72 texts were analyzed in total. To obtain a single score for all learners' texts, the total score obtained from the written texts was divided by 12. Researchers had ChatGPT write one text on the same topics and with the same guidelines. The data was collected during the 2023-2024 academic year.

Quantitative content analysis was applied to the texts. The texts were analyzed in terms of complexity, accuracy, fluency, and functional adequacy. The rubric developed by Kuiken and Vedder (2017) was used to determine functional adequacy. The rubric includes the dimensions of content, task requirements, comprehensibility, consistency, and coherence. Each dimension is scored on a scale of 1-6. The maximum total score that can be obtained is 24. Scoring was performed by two raters. Inter-rater reliability was examined using Kendal's Kappa analysis. The reliability coefficients were determined as content .771, task requirements .774, understandability .736, consistency and coherence .846, and the relationship between the scores was significant ($p < .01$). Complexity was analyzed at the sentence level. All sentences other than simple sentences were classified as complex sentences. Accuracy was analyzed at the word and sentence levels. Fluency was analyzed in terms of different word and word density. Functional adequacy was analyzed in terms of content, task requirements, comprehensibility, consistency, and coherence.

The analysis revealed that the average complex sentence ratio was 57.4% in learner texts at the A2 level and 55.3% in ChatGPT texts; 61.8% in learner texts at the B2 level and 65.1% in ChatGPT texts. The average correct sentence ratio was found to be 48.5% for A2 level learners and 100% for ChatGPT; 32.3% for B2 level learners and 98.6% for ChatGPT. The average correct word rate was found to be 92% for students at the A2 level and 100% for ChatGPT; 91.5% for students at the B2 level and 99.3% for ChatGPT. Comparisons were made in terms of different word usage and word density to determine fluency. Different word usage was 76.5% for A2 level students and 80.1% for ChatGPT; 65.9% for B2 level students and 65.8% for ChatGPT. Word density was 89.3% for A2 level students

and 88.1% for ChatGPT; 85.7% for B2 level students and 85.2% for ChatGPT. Functional adequacy scores were 17.8 for A2 level students and 22.3 for ChatGPT; 15.8 for B2 level students and 24 for ChatGPT.

In conclusion, this study demonstrates that ChatGPT has the potential to serve as a supportive tool in enhancing sentence and word-level accuracy and meaning construction during the process of developing writing skills in teaching Turkish as a foreign language, thereby ensuring the functional adequacy of the text. Functional adequacy encompasses presenting content that is sufficient in terms of meaning and relevant to the topic, meeting task requirements, writing texts with clearly understood purposes and ideas, and making the text a coherent whole with transition and connecting expressions. It appears that ChatGPT can be used to support learners of Turkish as a foreign language in organizing content that is grammatically correct and contextually appropriate. It would be more appropriate to use ChatGPT in the text planning or proofreading stages to help learners create original designs, preserve their creativity, and continue their linguistic development.