



# Türkçe Öğrenen Derlemi için Hata Etiketleme Sistemi

## An Error Coding System for the Turkish Learner Corpus

Anna GOLYNSKAIA\* 



### öz

Öğrenen derlemleri, bir dili yabancı veya ikinci dil olarak öğrenenlerin ortaya koyduğu metinlerin elektronik koleksiyonlarıdır. İkinci dil edinimi ve yabancı dil öğrenimi alanlarında yaygın olarak kullanılan öğrenen derlemleri, öğrenen dilini araştırmaya olanak sağlayan güvenilir araçlar olarak karşımıza çıkmaktadır. Bu çalışmanın amacı, 94 ülkeden gelen öğrencilere ait sınav kağıtlarından oluşan 274 bin kelimelik Türkçe Öğrenen Derleminin temelinde Türkçeye özgü bir hata etiketleme sistemini geliştirmektir. Derlemdeki metinler, orijinallerine sadık kalınarak manuel olarak bilgisayar ortamına aktarıldıktan sonra 44 bin kelimelik çekirdek derlem oluşturulmuş ve üzerinde hata etiketleme yapılmıştır. Hata etiketleme sistemi, ilk harfin hata alanını temsil ettiği, sonraki harflerin ise hata kategorisini ve geçerli ise kelime türünü belirttiği üç veya dört harfli kodlara dayanmaktadır. Toplamda 58 hata kodu mevcuttur. Tasarlanan hata etiketleme sistemi, Türkçe öğrenenlerin dil yeterliliğini değerlendirmek ve hata etiketlemesi içeren diğer Türkçe öğrenen derlemlerini oluşturmak için kullanılabilir. **Anahtar kelimeler:** Hata Etiketleme, Hata Taksonomisi, Hata Analizi, Öğrenen Derlemi, Yabancı Dil Olarak Türkçe

### ABSTRACT

Learner corpora are electronic collections of texts produced by learners of a foreign or second language. Learner corpora are reliable tools for investigating learner language and are widely used in the fields of second language acquisition and foreign language learning. This paper describes an error tagging system that has been designed on the basis of the 274,000-word Turkish learner corpus that comprises the Turkish examination papers written by learners coming from 94 different countries. The texts were manually keyboarded while retaining all errors, after which a 44,000-word component of the corpus was error coded using the specially devised error tag set. The majority of codes are based on a three- or four-letter system in which the first letter represents the error domain and the next series of letters identify the error category, as well as the word class where relevant. The error tag set has a total of 58 possible codes. The designed error tagging system can be used to assess the linguistic competence of Turkish learners and to build other error-annotated Turkish learner corpora. **Keywords:** Error Tagging, Error Taxonomy, Error Analysis, Learner Corpus, Turkish as a Foreign Language

\*(Dr. Öğr. Gör.), Sağlık Bilimleri Üniversitesi, Türk Dili Bölümü, İstanbul, Türkiye

ORCID: A.G. 0000-0001-6921-8110

#### Sorumlu yazar/Corresponding author:

Anna GOLYNSKAIA,  
Sağlık Bilimleri Üniversitesi, Türk Dili Bölümü,  
İstanbul, Türkiye  
E-posta: anna.golynskaia@sbu.edu.tr

**Başvuru/Submitted:** 08.12.2022

**Revizyon Talebi/Revision Requested:**  
26.12.2022

**Son Revizyon/Last Revision Received:**  
26.12.2022

**Kabul/Accepted:** 28.12.2022

**Online Yayın/Published Online:** 18.01.2023

**Atıf/Citation:** Golyńska, A. (2022). "Türkçe öğrenen derlemi için hata etiketleme sistemi". *Dilbilim Dergisi - Journal of Linguistics*, 39, 67-87. <https://doi.org/10.26650/jol.2022.1216218>



## EXTENDED ABSTRACT

Learner corpora, also known as interlanguage or L2 corpora, are defined as electronic collections of authentic foreign or second language data. Learner corpora differ from other data types used in the fields of second language acquisition (SLA) and foreign language teaching (FLT) in two main ways. Firstly, they are easily analyzable using various software tools due to being computerized. Secondly, learner corpora enable learner language to be correctly described and modeled due to the large amount of data they contain. In addition to the morphological and syntactic tagging and lemmatization commonly used in other corpus types, learner corpora are usually annotated with the help of a standardized system of error tags. During the error annotation process, errors are detected by one or more annotators who assign labels indicating the error type and add examples of the correct usage next to the labels. Error tagging is performed based on the error taxonomy developed for the given corpus. Error classifications used in learner corpora today can be divided into three groups: classifications indicating the source of error (e.g., morphology, lexis, case, number), classifications based on the ordering of the units in the source text (e.g., omission, addition, improper formation, improper ordering), and classifications that take various aspects into consideration, including error domain (e.g., morphology, lexicology, pragmatics), error category (e.g., diacritic usage, inflection, gender, mood), and word category.

This research discusses within its scope the process of designing an error annotation system based on the 44,000-word error-tagged component of a 274,000-word Turkish learner corpus that is comprised of Turkish examination papers written by learners from 94 different countries. The data were collected from eight Turkish teaching centers located in the Turkish provinces of Istanbul, Ankara, Sakarya, Samsun, Edirne, and Kocaeli that agreed to participate in the study. Therefore, the corpus comprises data produced by foreign students learning Turkish in an immersive language learning environment. Because the data collection phase coincided with the quarantine restrictions resulting from the COVID-19 pandemic, the study accessed the exam papers that had been stored in the archives of the Turkish teaching centers and scanned the sections on writing from the Turkish Proficiency Exam and C1 exams. The texts were manually keyboarded while retaining all errors. A tag was created containing metadata such as the nationality and gender of the student, the year of the exam, the subject, and the genre of writing was created for each text that had been saved as a separate Word file.

The digitization of the data that had been obtained from universities was completed in January 2021. In the process of transferring the texts to the computer environment, personal information (e.g., first and last name, university, residence address, age) was anonymized. Real information was replaced with randomly chosen names and numerical data; however, attention was paid to ensure that an equivalent word containing the same mistake was used in cases involving orthographic mistakes. In order to ensure that the maximum variety of texts

were included in the core corpus, a matrix was prepared that specified the students' genders, nationalities, writing topics, and genres. The core corpus contains 43,518 words and 50,487 tokens.

In addition to keywording the texts, the study investigated the error annotation systems developed for other corpora and identified the codes that could be used for the Turkish learner corpus. The preliminary error tags were divided into five categories: spelling, punctuation, morphology, syntax, and vocabulary. As the study progressed, it added other categories to these five, bringing the total number of error labels to 58. The majority of codes are based on a three- or four-letter system in which the first letter represents the error domain and the following letters identify the error category, as well as the word class where relevant.

The designed error tagging system can be used to assess Turkish learners' linguistic competence or to build other error-annotated Turkish learner corpora. In addition to the present study's contributions to the field, it also has some limitations. The research mostly focused on error labels related to the linguistic competence of Turkish learners, leaving discourse errors beyond the scope of the paper.

## 1. Giriş

Aradil veya hedef dil derlemleri olarak da adlandırılan öğrenen derlemleri, özgün yabancı veya ikinci dil verilerinin elektronik koleksiyonları olarak tanımlanmaktadır (Granger, 2008). Öğrenen derlemleri, ikinci dil edinimi (*SLA*) ve yabancı dil öğretimi (*FLT*) alanlarında yapılan araştırmaların kapsamında başvurulan diğer veri türlerinden başlıca iki açıdan farklılık göstermektedir. Bilgisayar ortamında hazırlandıkları için çeşitli yazılım araçları kullanılarak analiz edilebilmekte, ayrıca büyük miktarda veri içermeleri sayesinde öğrenen dilinin doğru ve kapsamlı bir şekilde tanımlanması ve modellenmesi için imkân yaratmaktadır (Granger 2013, s. 465).

Öğrenen derlemlerinde, diğer derlemlerde yaygın olarak kullanılan biçimbilimsel ve sözdizimsel etiketleme ile sözbirimleştirilmeye (*lemmatization*) ilaveten hata etiketlemesi yapılmaktadır. Etiketleme otomatik veya manuel olarak yapılabilir. Otomatik etiketleme özel yazılımlar aracılığıyla gerçekleştirilirken manuel hata etiketlemesi sürecinde bir ya da birkaç etiketleyici (*annotator*) tarafından kural dışı kullanım durumları tespit edilip hatalı kullanımlara hata türünün belirtildiği etiketler atanmakta ve yanlarına doğru kullanım örnekleri eklenmektedir (López, 2009). Hunston (2002), otomatik veya manuel etiketlemeye bağlı olarak elde edilen verilerin doğruluk oranlarının farklı olduğunun altını çizmektedir. Otomatik etiketleme söz konusu olduğunda derlemin tamamı büyüklüğüne bakılmaksızın etiketlenebilse de doğruluk oranının manuel olarak yapılan etiketlemeninkinden daha düşük olduğu dikkat çekmektedir.

Yöntem açısından değerlendirildiğinde, hata etiketlemesi kural tabanlı ve olasılık temelli olmak üzere iki şekilde yapılabilir. Kural tabanlı etiketleme dilbilgisi kuralları göz önünde bulundurularak yapılırken olasılık temelli etiketleme belli bir etiketin belirli bir bağlamda meydana gelmesinin istatistiksel olasılığına dayanmaktadır (Meyer 2002, s. 88).

Hata etiketleme sisteminin geliştirilmesi, hataların türlerine göre sınıflandırılmasını ve öğrenen derlemine yerleştirilecek hata etiketlerinin oluşturulmasını gerektirmektedir (Díaz-Negrillo ve Fernández-Domínguez 2006, s. 86). Hata etiketleme işleminden sonra, istatistiksel olarak anlamlı hataların elde edilmesi için istatistiksel analizler yapılmaktadır. Elde edilen bulgular, gelecekte yapılacak nitel araştırmaların odak noktasını belirlemeye ve/veya öğretim materyallerini yeniden yapılandırarak öğrenim sürecine müdahale etmeye imkân vermektedir (Díez Bedmar 2011, s. 36).

Günümüzde öğrenen derlemlerinde kullanılan hata sınıflandırmaları aşağıdaki başlıklar altında incelenebilmektedir:

a. Hatanın kaynağını gösteren sınıflandırmalar. Bu sınıflandırmalar biçimbilim, sözcükbilim ve söz dizimi gibi geniş dilbilimsel kategorilerin yanı sıra çatı, durum, sayı gibi daha spesifik kategorilere dayanabilmektedir.

b. Kaynak metindeki birimlerin sıralama şekline dayalı sınıflandırmalar: eksik kullanım (omission), fazla kullanım (addition), yanlış çekim veya türetim (misformation), yanlış sıralama (mis-ordering).

c. Hata alanı (biçimbilim, sözcükbilim, edimbilim), hata kategorisi (diyakritik kullanımı, çekim, cinsiyet, kip vb.) ve sözcük türü olmak üzere çeşitli boyutların hiyerarşik olarak ele alındığı sınıflandırmalar (Hana, Rosen, Škodová ve Štindlová 2010, s. 12).

Etkili olarak kullanılabilir hata etiketleme sisteminin sahip olması gereken özellikler şunlardır:

a. Bilgilendirici ve yönetilebilir olma: Hata etiketleme sistemi öğrenenlerin hataları hakkında yararlı bilgiler sağlamalı, ancak etiketleyicinin yönetemeyeceği kadar ayrıntılı olmamalıdır.

b. Yeniden kullanılabilir olma: Hata etiketleme sistemindeki kategoriler, çeşitli dillerde bulunabilecek nitelikte olmalıdır.

c. Esneklik: Etiketleme aşamasında etiketlerin eklenmesine veya silinmesine, etiketleme sonrası aşamada ise gerekli bilgilerin hızlı bir şekilde bulunmasına olanak sağlamalıdır.

d. Tutarlılık: Etiketleyiciler arasında tutarlılığın sağlanabilmesi için, hata kategorilerinin ayrıntılı açıklamalarıyla hata etiketleme ilkelerini içeren bir kılavuz hazırlanmalıdır (Granger 2013, s. 467).

Hata etiketlemesi yapılmış öğrenen derlemleri, etiketlenmemiş derlemlere göre birtakım avantajlara sahiptir. Etiketlenmiş derlemler, tüm verileri tarama ihtiyacını ortadan kaldırmaktadır. Bu tür derlemlerde (etiketlendiği takdirde) sadece doğru kullanımlar veya sadece yanlış kullanımlar görüntülenebilmektedir. Ayrıca derlem yazılımındaki araçlar kullanılarak öğrenenin yaşı, ana dili, dil seviyesi gibi bilgilere kolayca erişilebilmektedir. Son olarak etiketlenmiş öğrenen derleminde hata türüne göre arama yapılarak hataların sıklığıyla ilgili istatistik çıkarılabilmekte ve bu verilere dayanılarak hangi öğrenci grubunun hangi hatayı yapacağına dair tahmin yürütmek mümkün hale gelmektedir.

## 2. Türkçe Öğrenen Derlemi Hazırlama Süreci

Türkçe Öğrenen Derlemi, 2014-2020 yılları arasında Türkiye'deki devlet ve vakıf üniversitelerine bağlı Türkçe Öğretim Merkezlerinde (TÖMER'lerde) C1 Kur Sonu Sınavı ile Türkçe Yeterlik Sınavına giren 1245 yabancı uyruklu öğrencinin ortaya koyduğu 1515 yazılı anlatım örneğini (273.950 sözcük, 319.736 birim) içermektedir. Derlemin verileri Türkiye'nin İstanbul, Ankara, Sakarya, Samsun, Edirne ve Kocaeli illerinde bulunan ve çalışmaya katılmayı kabul eden 8 Türkçe Öğretim Merkezinden toplanmıştır. Böylelikle derlemimizi Türkçeyi hedef dil ortamında öğrenen yabancı öğrencilerin yazılı metinleri oluşturmuştur. Türkçeyi yurtdışındaki merkezlerde ana dili/ikinci dil ortamında öğrenenler çalışmamızın dışında tutulmuştur. Veri toplama aşaması COVID-19 salgınından dolayı uygulanan karantina ve sokağa çıkma yasağı gibi kısıtlamalara denk geldiği için, TÖMER'lerin arşivlerinde saklanan sınav kağıtlarına erişilerek yazma bölümleri taranmıştır. Sınav kâğıtları elde edilen çoğu öğrencinin yaş ve ana dili bilgilerine ulaşılamamıştır. Bunun nedeni, kursiyerlerin doğum tarihlerinin/yaşlarının ve ana dillerinin öğrenci listelerinde belirtilmemiş olup bu tür bilgilerin, erişilmesine izin verilmeyen pasaport fotokopilerinde bulunmasıdır.

Veri toplama işlemi devam ederken eş zamanlı olarak elde edilen metinler asıllarına sadık kalınarak bilgisayar ortamına aktarılmıştır. Aynı bir Word dosyası olarak kaydedilen her metnin, sınava giren öğrencinin uyruğu ve cinsiyeti gibi demografik bilgilerle sınav yılı, yazma konusu ve türü (kompozisyon, e-posta, dilekçe vs.) gibi bilgileri içeren künyesi oluşturulmuştur.

Üniversitelerden elde edilen verilerin dijitalleştirilmesi 2021 yılının ocak ayında tamamlanmıştır. Metinlerin bilgisayar ortamına aktarılması sürecinde öğrencilere ait kişisel verilerin gizliliği göz önünde bulundurularak ad, soyadı, okuduğu üniversite, ikamet adresi, yaş vb. kişisel bilgiler anonim hâle getirilmiştir. Gerçek bilgiler rastgele seçilmiş isim ve sayısal verilerle değiştirilmiş, ancak kişisel verilerin yazılışında bir hata yapıldıysa muadil kelimenin aynı hatanın oluşmasını sağlayan bir kelime olmasına dikkat edilmiştir.

### 3. Çekirdek Derlemin Oluşturulması

Ana derlem oluşturulduktan sonra hata etiketlemesi yapılabilmesi için ana derlemi temsil gücüne sahip çekirdek derleme dâhil edilecek metinler belirlenmiştir. Çekirdek derlemi oluşturma ihtiyacını, hata etiketlemesini tek kişinin manuel olarak yapması doğurmuştur.

Hedef evreni temsil edebilecek çekirdek derlemin kaç yazıyı içermesi gerektiğini tespit etmek için örneklem büyüklüğü hesaplayıcısı kullanılmıştır. Hesaplamaların sonucu olarak minimum örneklem büyüklüğünün 307 olması gerektiği saptanmıştır (Güven düzeyi: %95, güven aralığı: 5). Derleme dâhil edilecek metinlerin maksimum çeşitliliğini sağlamak amacıyla öğrencilerin cinsiyetleri, uyrukları, yazma konuları ve metin türlerinin belirtildiği bir matris hazırlanmıştır. Metinler kadın erkek sayısı gözetilerek seçilmiştir. Çekirdek derlemi oluşturan metinlerin yarısı kadın öğrenciler, diğer yarısı ise erkek öğrenciler tarafından yazılmıştır. Çekirdek derleme, ulaşılabilen tüm ülkelerin (94 ülkenin) temsilcilerinin yazdığı metinler dâhil edilmiştir. Yazıların konu dağılımına bakıldığında Tolstoy'un "Acı duyabiliyorsan canlısın. Başkalarının acısını duyabiliyorsan insansın." sözleriyle ilgili yazılarla "Eğitim" konulu kompozisyonların sayıca daha fazla olduğu görülmektedir. Öğrencilerin ortaya koyduğu yazılar metin türü bakımından incelendiğinde kompozisyon türündeki metinlerin diğer metin türlerine göre daha fazla olduğu ortaya çıkmaktadır. Çekirdek derlemin boyutu değerlendirildiğinde sözcük sayısının 43.518, birim sayısının ise 50.487 olduğu görülmektedir.

### 4. Hata Etiketleme Sistemi

Türkçe Öğrenen Derleminde kullanılacak etiketleri tespit etmek için çeşitli derlemler için geliştirilen hata etiketleme sistemleri (Dagneaux, Denness, Granger, Meunier, Neff ve Thewissen, 2005; Izumi, Uchimoto ve Isahara, 2005; Nicholls, 2003) incelenmiş ve Türkçenin yapısına uygun bir sistem geliştirilmeye çalışılmıştır. Bu amaçla 25 metinden oluşan bir pilot çalışma uygulanmıştır. İlk aşamada öğrencilerin yaptığı hatalar yazım, noktalama, biçimbilim, söz dizimi ve söz varlığı olmak üzere beş başlık altında toplanmış ve bu başlıklara uygun hata etiketleri hazırlanmıştır. Çalışma ilerledikçe yukarıda bahsi geçen beş kategoriye başka kategoriler de eklenmiş ve hata etiketlerinin toplam sayısı 58'e ulaşmıştır (Bkz. Tablo 1).

Bu noktada, çekim ve yapım ekleri açısından zenginlik gösteren ve bilgi yapısı güdümlü söz dizimine sahip olan Türkçeyi yabancı dil olarak öğrenenlerin yaptığı hataları etiketlemenin zorlayıcı bir süreç olduğunun belirtilmesi gerekmektedir. Önerilen hata etiketleme sistemi, Türkçe gibi sondan eklemeli dillerde yapılan hatalı kullanımları etiklerken etiketleme sürecinin sınırlamaları ile derlem kullanıcılarının ihtiyaçları arasında bir uzlaşma sağlama girişimidir.

<b>Tablo 1: Hata Etiketleri</b>		
<b>Hata Alanı</b>	<b>Hata Etiketi</b>	<b>Etiketin Açılımı</b>
Noktalama Hataları (N)	EN YN	Eksik Noktalama Yanlış Noktalama
Yazım Hataları (Y)	Y YBA YD YBH YUSY YUU YUDS YKH YUB YUD YUT YKS YUST YKB YE	Genel Yazım Hataları Yazım Bitişik Ayrı Yazım Diyakritik Yazım Büyük Harf Yazım Ünsüz Yumuşaması Yazım Ünlü Uyumu Yazım Ünlü Düşmesi Yazım Kaynaştırma Harfi Yazım Ünsüz Benzeşmesi Yazım Ünlü Daralması Yazım Ünlü Türemesi Yazım Kısaltma Yazım Ünsüz Türemesi Kelimelerin Bölünmesi Eklerin Yanlış Yazımı
Biçimbilimsel Hatalar (B)	BID BIS BII BIT	İsim Durum İsim Sayı İsim İyelik İsim Tamlayan
	BFC BFZ BFSY BFO BFS BFK BFY BFCM BFB BEF	Fiil Çatı Fiil Zaman Fiil Sayı Fiil Olumsuzluk Fiil Şahıs Fiil Kip Yeterlilik Fiili Çekimli/Çekimsiz Fiiller Birleşik Fiil Ek-Fiil
	BIF BSF BZF	İsim-fiil Sıfat-fiil Zarf-fiil
	BZMS BZMD BDZM BZMT BZMI BILZM	Zamir Sayı Zamir Durum Dönüşlülük Zamiri Zamir Tamlayan Zamir İyelik İlgi Zamiri

Biçimbilimsel Hatalar (B)	BIL	İlgeç
	BKT	Kelime Türü
	BGE	Gereksiz Ek
	BSE	Soru Eki
	BSRE	Şart Eki
Sözcüksel Hatalar (S)	SAK SYK SEK SGK	Anlaşılmayan Kelime(ler) Yanlış Kelime(ler) Eksik Kelime(ler) Gereksiz Kelime(ler)
Sözdizimsel Hatalar (SD)	SDKS SDEC SDN SDDA	Kelime Sırası Eksik Cümle Nesne Dolaylı Anlatım
Anlamsal Hatalar (A)	AAB AIH AACs	Anlam Belirsizliği İfade Hatası Anlamca Çelişen Sözcükler
Metin Biçimlendirme Hataları (MB)	MBB MBS	Başlık Satır

## 5. Hata Etiketlerinin Kullanım Alanları

### 5.1. Noktalama Hataları

“EN” (eksik noktalama) ve “YN” (yanlış noktalama) şeklinde kodlanan hatalar noktalama işaretlerinin kullanılmaması veya yanlış kullanılması ile ilgilidir.

İnsan dünyada tek başına yaşamıyor <err type=»EN»> </err> <corr type=»EN»>. </corr>

Çünkü <err type=»YN»>, </err> <corr type=»YN»> </corr> arkadaş seçimleri bu hayattaki en önemli seçimlerdir.

### 5.2. Yazım Hataları

#### 5.2.1. Genel Yazım Hataları

“Y” olarak kodlanan hata türü genel yazım yanlışlarını ifade etmektedir. Sıklıkla yapılan yazım hataları; harf sırasının karıştırılması, harf düşmesi, harf türemesi, ünsüz ikizleşmesi, kelimelerin yabancı karakterler kullanılarak veya yabancı dilde telaffuz edildiği gibi yazılması, *e* ve *i* harflerinin karıştırılması, *a* ve *e* harflerinin karıştırılması ve *ğ* ve *y* harflerinin karıştırılmasıdır.

<err> tecürbe </err><corr> tecrübe </corr> <err>

<err> büyüdük </err><corr> büyüdüük </corr>

<err> miğdemiz </err><corr> midemiz </corr>

<err> bittirdi </err><corr> bitirdi </corr>

<err> stress </err><corr> stres </corr>



<err> tehlikeli </err><corr> tehlikeli </corr>  
 <err> hadiye </err><corr> hediye </corr>  
 <err> geleceğini </err><corr> geleceğini </corr>

### 5.2.2. Kelimelerin Bitişik/Ayrı Yazımları ile İlgili Hatalar

Bu kategorideki hatalar birkaç başlık altında incelenebilmektedir:

a) Da/de bağlacının yazımıyla ilgili hatalar:

Bir insanın kişiliğinin oluşmasında akrabaların, <err> arkadaşlarında </err><corr> arkadaşların da </corr> çok etkisi var bence.

b) Yeterlik fiilinin yazımıyla ilgili hatalar:

... istediğimiz şeyin gerçekleşmemesine yol <err> açabilir </err><corr> açabilir </corr>.

c) Birbiri, birçok, birkaç, bugün, herhangi, herkes, neyse, veya gibi bitişik yazılması gereken kelimelerle ilgili hatalar:

Bence <err> her kes </err><corr> herkes </corr> bir kere olsa da kendi şansını denemeli...

d) Bir şey, her şey, bir arada/bir araya, güler yüzlü, şu an, ya da, yanı sıra gibi ayrı yazılması gereken kelimelerle ilgili hatalar:

<err> Şuan </err><corr> Şu an </corr> bir amacım bile yok.

e) Etmek, olmak yardımcı fiilleriyle oluşturulan birleşik fiillerle ilgili hatalar:

İşte insanlar sürekli bu yüzden <err> kayb ediyorlar </err><corr> kaybediyorlar </corr>.

f) Ki bağlacının yazımıyla ilgili hatalar:

Tolstoy <err> diyorki </err><corr> diyor ki </corr>: ...

### 5.2.3. Diyakritik Hataları

“YD” etiketi diyakritik işaretlerle ilgili hataları ifade etmektedir. Bu hatalar tek nokta imi, çift nokta imi, düzeltme imi, çengel imi veya yumuşatma iminin kullanılmamasından veya tam tersine gereksiz kullanılmasından kaynaklanmaktadır.

<err> basit </err><corr> basit </corr>

<err> boyle </err><corr> böyle </corr>

<err> âksi </err><corr> aksi </corr>

<err> arac </err><corr> araç </corr>

<err> götürdü </err><corr> götürdü </corr>

### 5.2.4. Büyük Harflerin Kullanımıyla İlgili Hatalar

“YBH” ile kodlanan hata türü sıklıkla başlıklarda ve cümle başındaki kelimelerle özel isimlerin yazılışında görülmektedir.

Yurt <err> dışında yaşama </err><corr> Dışında Yaşama </corr>

Sevgili <err> sevda </err><corr> Sevda </corr>, canım arkadaşım.

<err> insan </err><corr> İnsan </corr> hayatında <err> Aile </err><corr> aile </corr> ve <err> Aile </err><corr> aile </corr> ortamı çok önemli...

### 5.2.5. Türkçenin Ses Yapısıyla İlgili Hatalar

Derlemdeki yazım hatalarının bir bölümü, Türkçenin ses yapısı, ses olayları ve onların yazı diline yansımaları açısından ele alınmıştır.

- a) Ünlü uyumu ile ilgili hatalar (YUU): <err> hastaneya </err><corr> hastaneye </corr>
- b) Ünsüz Yumuşaması ile İlgili Hatalar (YUSY): <err> hissetiyorsan </err><corr> hissediyorsan </corr>
- c) Ünlü düşmesi ile ilgili hatalar (YUDS): <err> kaydı </err><corr> kaydı </corr>
- d) Ünsüz benzeşmesi ile ilgili hatalar (YUB): <err> kötülükden </err><corr> kötülükten </corr>
- e) Ünlü daralması ile ilgili hatalar (YUD): <err> isteyorum </err><corr> istiyorum </corr>
- f) Ünsüz türemesi ile ilgili hatalar (YUST): <err> hakım </err><corr> hakkım </corr>
- g) Satır sonunda kelimelerin bölünmesi ile ilgili hatalar (YKB): <err> edin-iz </err><corr> edi-niz</corr>

### 5.2.6. Kaynaştırma Harfleri ile İlgili Hatalar

Kaynaştırma harflerinin kullanımıyla ilgili hatalar derlemede “YKH” kodu ile etiketlenmiştir. Bu tür hatalar iyelik eklerinden sonra “n” yerine “y” kaynaştırma harfinin kullanılması veya ek-fiilleri ünlü ile biten bir kelimeyle birleştirdiklerinde araya “y” harfinin konmaması sonucu ortaya çıkmaktadır.

- <err> aşklarından </err><corr> aşklarından </corr>
- <err> iyimiş </err><corr> iyiymiş </corr>

### 5.2.7. Eklerin Yazımı ile İlgili Hatalar

Hata etiketleme sürecinde bazı eklerin yazılışında düzenli olarak hata yapıldığı tespit edildiğinden eklerin yazımıyla ilgili hatalar ayrı bir kategoride değerlendirilerek “YE” koduyla etiketlenmiştir. Yazımında en çok hata yapılan ekler; -Dik eki, geniş zaman eki, gelecek zaman eki, tamlayan eki, -An eki ve -IncA ekidir.

- <err> büyüdünde </err><corr> büyüdüğünde </corr>
- <err> düzelter </err><corr> düzeltir </corr>
- <err> harcayacayım </err><corr> harcayacağım </corr>
- <err> insannın </err><corr> insanın </corr>
- <err> yaşan </err><corr> yaşayan </corr>
- <err> duydunca </err><corr> duyunca </corr>

### 5.2.8. Kısaltmaların Yazımı ile İlgili Hatalar

Günümüzde kısaltmalar özellikle dijital ortamın etkisiyle yaygınlık kazanmıştır. Öyle ki Türkçe öğrenen yabancı öğrencilerin sınav kâğıtlarında da Türkçe Yazım Kılavuzunda yer almayan kısaltmalara rastlanmıştır. Bunun dışında öğrencilerin yazılarında vs. gibi daha

geleneksel kısaltmaların bulunduğu görülmektedir. Derlem etiketlemede “YKS” kodu yabancı öğrencilerin kısaltmalarda yaptığı hataları ifade etmek için kullanılmıştır.

Plaja gidebilmek için de en az 15 <err> dk </err><corr> dakika </corr> yürümemiz gerekiyordu.

<err> vesa </err><corr> vs. </corr>

### 5.3. Biçimbilimsel Hatalar

Biçimbilimsel hatalar, farklı kelime türlerinin dilbilgisel kategorileri (sayı, durum, iyelik, zaman vb.) ve eklerin kullanımı ile ilgilidir.

#### 5.3.1. Hâl Eklerinin Kullanımıyla İlgili Hatalar

Hâl eklerinin kullanımıyla ilgili hatalar Türkçe öğrenenlerin yaptığı biçimbilimsel hataların başında gelmektedir. “BİD” etiketi, öğrenenlerin durum eklerinin kullanım yerlerini, işlevlerini ve kurdukları anlamsal ilişkileri tam olarak kavrayamadıkları için uygulama aşamasında yeterli başarıyı gösterememeleri sonucu ortaya çıkan hataları ifade etmektedir.

Çünkü <err> dünyada </err> <corr> dünya </corr> biraz karışıktır <err/>...

İnsanlar bütün <err> işleri </err> <corr> işlere </corr> başlarken çiraklıktan başlar.

<err> Sabah </err> <corr> Sabahtan </corr> akşama kadar...

... yani insanlar şahane bir <err> dünya </err> <corr> dünyada </corr> yaşamak istemektedir.

#### 5.3.2. Tamlayan Ekinin Kullanımı ile İlgili Hatalar

“BİT” olarak etiketlenen hataların büyük çoğunluğu, belirtili tamlamalarda veya “-mA” ve “-DİK” ile kurulan yan cümlelerde tamlayan ekinin kullanılmamasından ortaya çıkmaktadır.

Televizyon günümüz şartlarında <err> hayatımız </err> <corr> hayatımızın </corr> bir parçası hâline geldi.

... bu <err> kelime </err> <corr> kelimenin </corr> ne demek olduğu bile önemli değil, diye düşünüyoruz.

#### 5.3.3. İyelik Eklerinin Kullanımı ile İlgili Hatalar

“Bİİ” olarak kodlanan hatalar, isimlerin kime veya neye ait olduklarının belirtilmemesinden, isimlere gereksiz yere iyelik eklerinin getirilmesinden veya iyelik eklerinin yanlış kişi ve sayı bilgisini kodlamasından kaynaklanan hatalardır.

Lakin okumalarına rağmen <err> ülkene </err><corr> ülkelerine </corr> katkı sağlamıyorlar <err/>...

Bilim <err> adamlara </err><corr> adamlarına </corr> göre...

Mesela, benim bir <err> arkadaş </err><corr> arkadaşım </corr> kitap okumayı çok seviyor.

### 5.3.4. Çoğul Ekinin Kullanımı ile İlgili Hatalar

“BIS” olarak kodlanan hatalar, Türkçe öğrenenlerin ana dillerinden ya da bildikleri diğer yabancı dillerden edindikleri alışkanlık gereği, çoğul olarak kullanılmayan veya cümlenin gidişatına göre çoğul eki alması gerekmeyen isimler ile sayılar ve “birkaç, az, biraz, çok, birçok, çoğu, her, herhangi bir, hiç, hiçbir, bir sürü, çok sayıda” gibi anlamında çoğulluk bulunan ifadelerden sonra gelen isimlere çoğul ekini getirmelerinden veya tam tersine çokluk anlamı taşıyan isimleri kullanmaları gerektiren durumlarda çoğul ekini getirmeyi unutmalarından kaynaklanmaktadır.

İnsanlar çok fazla sosyal <err> medyalar </err><corr> medya </corr> kullanıyor.

... hayatın her <err> alanlarında </err><corr> alanında </corr>...

Binlerce <err> çocuklar </err><corr> çocuk </corr> var.

Yani çeşitli <err> araştırma </err><corr> araştırmalar </corr> yapabiliriz.

### 5.3.5. Zaman Eklerinin Kullanımı ile İlgili Hatalar

Türkçe öğrenenlerin yazılı anlatım örneklerindeki “BFZ” olarak kodlanan hataların bir kısmı, zaman eklerinin kullanımındaki tutarsızlıktan kaynaklanmaktadır. Öğrenenler anlatımlarını sürdürürken yanlış olarak bir zaman düzleminden başka birine atlayıp paragrafın akışını bozabilmektedir.

... suç işleyen insanlar </corr><err> insanlarını </err><corr> insaları </corr> öldürür ve eşyaları <err> alıyor </err><corr> çalar </corr>.

Bunun dışında “BFZ”, Türkçe öğrenenlerin çeşitli zamanlar arasındaki farkı kavrayamamaları sonucu ortaya çıkan hataları etiketlemek için kullanılmaktadır.

Birkaç saat <err> bekliyorduk </err><corr> bekledik </corr>.

Eğer ben de iyi çalışsam, iyi puan alsam nasıl <err> oluyor </err><corr> olur </corr> <corr>?

Ama bir kişi, bu yöntemleri birlikte kullanırsa daha bilgili <err> olacak </err><corr> olur </corr>.

### 5.3.6. Fiil Çatılarının Kullanımı ile İlgili Hatalar

Etiketi “BFC” olan hataların sebebi, öğrenenlerin kurduğu cümlelerdeki çatı uyumsuzluğudur.

... şüpheli insanlar tarafından <err> kullanabilir </err><corr> kullanılabilir </corr>.

Yabancı bir dil öğrenirken zorluklarla <err> karşılayabiliriz </err><corr> karşılaşılabiriz </corr>.

Şu anda Tömer <err> bitirmek </err><corr> bitmek </corr> üzere.

... ama sevgi zamanla <err> besler </err><corr> beslenir </corr>.

### 5.3.7. Kiplerin Kullanımı ile İlgili Hatalar

“BFK” olarak kodlanan hatalar çoğunlukla Türkçe öğrenenlerin şimdiki zaman ya da geniş zaman 2. tekil kişi ile çekimlenmiş fiiller ile şart kipi 2. tekil kişi ile çekimlenmiş fiilleri karıştırmaları sonucu ortaya çıkmaktadır.

Eğer </corr> kocan fakir olursa onunla yaşamaya razı <err> olursan </err><corr> olursun </corr>, çünkü onu <err> seviyorsan </err><corr> seviyorsun </corr> ve başkasını <err> istemiyorsan </err><corr> istemiyorsun </corr>.

### 5.3.8. Çekimli ve Çekimsiz Fiillerin Kullanımı ile İlgili Hatalar

Etiketi “BFCM” olan hatalı kullanımlarda çekimsiz fiiller çekimli fiillerin işlevini görmekte veya tam tersine çekimli fiiller isim ve zarf görevinde kullanılmaktadır.

İnsan başka insanları <err> hissetmek </err><corr> hissetmeli </corr>.

En önemli sorun çok para <err> gerekiyor </err><corr> gerekmesi </corr>.

### 5.3.9. Fiillerde Şahıs Eklerinin Kullanımı ile İlgili Hatalar

“BFS” olarak etiketlenen hatalar öznenin bildirdiği şahısla yüklem bildirdiği şahıs arasında uyumun bulunmamasından kaynaklanmaktadır.

... bu konu </corr> aklıma <err> aklıma yattım </err><corr> aklıma yattı </corr>.

... herkes <err> çalışmalıyız </err><corr> çalışmalı </corr>...

### 5.3.10. Olumsuzluk Ekinin Kullanımı ile İlgili Hatalar

“BFO” şeklinde kodlanan hatalar cümlenin anlamı gereği olumlu olması gereken çekimli ve çekimsiz fiillere olumsuzluk ekinin getirilmesi veya tam tersine olumsuzluk kavramı veren fiil ve fiilimsilerde -ma/-me ekinin kullanılmaması sonucu ortaya çıkmaktadır.

Bu kısa hayatımızı pişmanlıkla ve üzülerek <err> geçirmemeyelim </err><corr> geçirmeyelim </corr>.

...<corr> hiçbir </corr> zaman <err> unutacağım </err><corr> unutmayacağım </corr>.

### 5.3.11. Birleşik Fiillerin Kullanımı ile İlgili Hatalar

“BFB” olarak kodlanan hatalar, yeterli fiilin kullanılmasının uygun olabileceği cümlelerde yeterlik fiili kullanılmadığında veya sonu ünlü harfle biten fiillerden sonra yeterlik fiili “Abil...” yerine sadece “bil...” biçiminde getirildiğinde meydana gelmektedir.

... daha doğrusu ne yapabilirim? </s><s> Zamanı geri <err> çevirmem </err><corr> çeviremem </corr>...

<err> harcabilirsin </err><corr> harcayabilirsin </corr>

### 5.3.12. Ek-Fiilin Kullanımı ile İlgili Hatalar

Etiketi “BEF” olan hataların sebebi çoğunlukla ek-fiilin gereksiz kullanılması veya hiç kullanılmamasıdır.

Canlılar <err> gibidir </err> <corr> gibi </corr> rutin bir hayat sürüyor...

<err> Aralık'ta </err> Aralıkta benim Türkçem iyi <err> değil </err> <corr> değildi </corr> ...

### 5.3.13. İsim-Fiillerin Kullanımıyla İlgili Hatalar

“BIF” olarak kodlanan hataların çoğu, -mA ve -mAk ekiyle türetilen isim-fiillerin kullanım alanlarının karıştırılması sonucu doğmaktadır.

...bu sebeple <err> boşanmak </err> <corr> boşanma </corr> oranı çok yükseldi.

Her şey bir gün son bulacak bu sonların güzel <err> olmak </err> <corr> olması </corr> yalnızca insanların elindedir.

... her şeyimi <err> toplamaya </err> <corr> toplamak </corr> zorundaydım.

Aslında hayal, hayatımızı devam <err> ettirmek </err> <corr> ettirmemiz </corr> için bize enerji veriyor.

### 5.3.14. Sıfat-Fiillerin Kullanımıyla İlgili Hatalar

“BSF” şeklinde kodlanan hatalar, -An ekiyle yapılan sıfat-fiillerin -Dik ekiyle yapılan sıfat-fiillerle karıştırılması veya isim-fiillerin girişik birleşik cümlelerin yan cümlesinde -Dik/-AcAk ekleriyle yapılan sıfat-fiil görevinde kullanılması sonucu ortaya çıkan hatalardır.

Dünyanın en zengin kişilerinden birisi, Microsoft şirketini <err> kurduğu </err> <corr> kuran </corr> Bill Gates dedi ki...

... aile <err> bireyleri </err> bireyelerinin çocuklarına hangi değerleri <err> aşılması </err> <corr> aşladığı </corr> önemlidir.

### 5.3.15. Zarf-Fiillerin Kullanımıyla İlgili Hatalar

“BZF” etiketi, öğrencilerin yanlış zarf-fiil eklerini getirmelerinden veya zarf-fiilleri kullanmaları gereken cümlelerde çekimli fiilleri, isim-fiilleri, sıfat-fiilleri ve dil bilgisel olmayan birtakım yapıları kullanmalarından kaynaklanan hataları ifade etmektedir.

Arabaları <err> kullanınca </err><corr> kullanırken </corr> özen göstermemiz, dikkat etmemiz lazım...

<err> Çalışmam için </err> <corr> Çalıştığım için </corr> sınavda başarılı oldum ve ondan sonra burs kazandım.

<err> Söylediklerini gibi </err> <corr> söyledikleri gibi </corr> “keskin sirke küpüne zarardır” <err/>.

### 5.3.16. Zamirlerin Kullanımıyla İlgili Hatalar

Zamirlerin kullanımıyla ilgili hatalar; zamirlerin tekillik-çoğulluk durumuyla ilgili hatalar (BZMS), zamirlerin hâl ekleriyle kullanımında yapılan hatalar (BZMD), zamirlerin iyelik ekleriyle kullanımında yapılan hatalar (BZMI), zamirlerin tamlayan ekiyle kullanımında yapılan hatalar (BZMT), dönüşlülük zamirinin kullanımıyla ilgili hatalar (DZM) ve ilgi zamirinin kullanımıyla ilgili hatalar (BILZM) olmak üzere altı grupta incelenmiştir.

Mağazalarda çok az <err> şeyler </err> <corr> şey </corr> <err> vardı.

Bazı insanlar için para mutluluk getirebilir, bazıları bazılarına ise getirmeyebilir.

Dünyanın eğitim sistemi çok kötü olduğu için seçtiğiniz bölümde kendi kendine kendi kendinize çalışmalısınız.

Neyse ki benim kötü bir kişi olduğumu bazıları biliyorlar.

Eğer bir kişi onun kendi işini sevmiyorsa...

Bizimki Bizim eğitim sistemimiz, Türkiye'dekinden biraz farklıdır.

### 5.3.17. İlgeçlerin Kullanımıyla İlgili Hatalar

“BIL” etiketi ilgeçlerin yerine hâl eklerinin, hâl eklerinin yerine ise ilgeçlerin kullanıldığı hataları belirtmektedir.

...çünkü onlar çocuklarını koruyup kolluyorlar ve her şeyin en iyisini çocuklarına çocukları için istiyorlar.

...bazen çok merak ediyorum bu hakkında bunu...

### 5.3.18. Kelime Türleri ile İlgili Hatalar

“BKT” olarak kodlanan hatalar Türkçe öğrenenlerin, çoğunlukla sıfatlarla isimleri birbirinden ayırt edememeleri sonucu ortaya çıkmaktadır.

... insan yorucudan yorgunluktan ölüyor.

Bunun için çok dikkatli dikkatli olmalıyız.

İkinci faydalı faydası şu...

### 5.3.19. Yapım Eklerinin Gereksiz Kullanımı ile İlgili Hatalar

“BGE” etiketi, bir kelimeye gereksiz yere bir yapım eki getirilmesi sonucu yanlış bir türevin ortaya çıktığı durumlarda kullanılmaktadır.

Önce kendimi tanıtıracam tanıtacağım...

... evde kalmak yurttan kalmaktan daha rahat rahat.

... huzurlu, mutlulu mutlu yaşaması her insanın hakkıdır.

... bazıları yoksulluklara yoksullara iyi davranmıyor.

### 5.3.20. Soru Ekinin Kullanımı ile İlgili Hatalar

“BSE” etiketinin kullanıldığı durumlar aşağıdaki gibidir:

a. Evet/Hayır cevabını gerektiren soru cümlelerinde soru ekinin kullanılmaması

... yoksa ondan daha önemli işlerimiz var mı?

b. Soru zamirleriyle soru zarflarını içeren cümlelerde soru eki kullanılması

... parayı ne zaman harcayabilir miyiz? harcayabiliriz?

### 5.3.21. Şart Ekinin Kullanımı ile İlgili Hatalar

“BSRE” etiketi, şart ekinin getirilmediği ya da yanlış zaman ekiyle beraber kullanıldığı durumlarda kullanılmaktadır.

Eğer bir kişi <err> hasta </err><corr> hastaysa </corr>...

Üstelik bu üniversiteden mezun <err> olacaksan </err><corr> olursan </corr> <err> çalışmak için yurt dışına gidebilirsin.

## 5.4. Sözcüksel Hatalar

### 5.4.1. “Anlaşılmayan Kelime” Olarak Kodlanan Hatalar

“SAK”, öğrenenlerin Türkçede mevcut olmayan kelimeleri kullanmalarından ya da Türkçe kelimeleri kullanmalarına rağmen bu kelimeleri kullanma amaçlarının anlaşılmasından kaynaklanan hataları etiketlemek için kullanılmaktadır.

Sosyal medya hastalıklara neden olur. Bütün <err> noktalar </err><corr> ? </corr> iyi ve doğru kullanmamız lazım.

<err> Hayi </err><corr> ? </corr> <err> grub </err><corr> grup </corr> iş birliği göstermiyorsa ceza vermeli.

### 5.4.2. “Yanlış Kelime” Olarak Kodlanan Hatalar

“SYK” etiketi çoğunlukla kelimelerin ses benzerliği nedeniyle karıştırılması sonucu ortaya çıkan hataları, eşdizim hatalarını ve yakın anlamlı veya aynı semantik alana ait kelimelerin yanlış kullanımıyla ilgili hataları ifade etmektedir.

İlk olarak her insanın bir <err> çevirisi </err> <corr> çevresi </corr> olması gerekir...

Orada kısa süre çalışacağım ve tecrübe de <err> alacağım </err> <corr> kazanacağım </corr>.

İyisini yaparken <err> rahatlamak </err> <corr> rehavete kapılmak </corr> yerine daha iyisini yapmak için elimizden ne geliyorsa yapmalıyız.

### 5.4.3. “Eksik Kelime” Olarak Kodlanan Hatalar

“SEK”, bir cümle ögesinin eksik olduğu kullanımları veya çoğunlukla görevli kelimelerin kullanılmamasından kaynaklanan hataları etiketlemek için kullanılmaktadır.

... savaşlarda binlerce <corr> insan </corr> birbirini öldürüyor...

Mesela senin bilmediklerini onlar <corr> biliyor </corr>, onların bilmediklerini sen biliyorsun.

... ve önemli <corr> bir </corr> şey yapabilirim.

... ve onları nasıl koruyabilirim <corr> diye </corr> düşünmeliyiz.

### 5.4.4. “Gereksiz Kelime” Olarak Kodlanan Hatalar

“SGK” şeklinde kodlanan hataları genellikle bağlaçların, zamirlerin, ilgeçlerin ve “bir” kelimesinin gereksiz kullanılması sonucu ortaya çıkan hatalar teşkil etmektedir.

Ayrıca sevdiğiniz insanla yaşayabilirsiniz <err> ve </err> fakir olsa da...

Bazı öğrenciler <err> onlar </err> <err> akabalarıyla </err> akrabalarıyla geliyordu...



Çünkü yaşamak <err> için </err> şimdi çok tehlikeli...  
Güzel <err> bir </err> sokaklarda insanların gezmesi...

## 5.5. Sözdizimsel Hatalar

### 5.5.1. Kelime Sırası Hataları

“SDKS” olarak etiketlenen hatalar, Türkçe öğrenenlerin ana dillerinden Türkçeye doğrudan çeviri mantığıyla düşündükleri için kelime sırası açısından bozuk cümleler kurduklarında ortaya çıkmaktadır.

<err> Bir fazla kilolu </err><corr> Fazla kilolu bir </corr> kız vardı.

<err> İkinci söylemek istediğim </err><corr> Söylemek istediğim ikinci </corr> şey şudur...

<err> Nasıl bunu </err><corr> Bunu nasıl </corr> yapabilirsiniz?

### 5.5.2. Dolaylı Anlatım Hataları

“SDDA” etiketi Türkçe öğrenenlerin dolaylı anlatımlı cümleleri kuramamalarından kaynaklanan hataları belirtmektedir.

<err> Onunları </err><corr> Onlar </corr> <err> bilmiyorlar gerçek mutluluk ne demek </err><corr> gerçek mutluluğun ne demek olduğunu bilmiyorlar </corr>.

İspanyol Konsolosluğu'na gitmek <err> zorundayıma </err><corr> zorunda olduğumu </corr> biliyordum.

### 5.5.3. Cümlelerin Önemli Bir Kısmının Eksik Bırakılmasından Kaynaklanan Hatalar

“SDEC” şeklinde kodlanan hatalar Türkçe öğrenenlerin kurduğu cümlelerde bir ya da birden çok ögenin eksik bırakılması sonucu ortaya çıkan anlatım bozukluklarını veya yazı diline has olmayan kullanımları ifade etmektedir.

Bu <err> müdüne </err><corr> müdür </corr> her <err> şey </err><corr> şeyi </corr> anladı, ama <err> fikir yok </err><corr> durumu nasıl düzeltereğine dair hiçbir fikri yoktu </corr>.

Geçmişte insanlar bir evi alırken komşuların durumuna ve davranışlarına bakıyordu çünkü ev alma komşu al <corr> diye bir söz var </corr>.

### 5.5.4. Nesne Hataları

Türkçe öğrenen yabancı öğrencilerin belirtili nesnelere belirtisiz nesnelere arasındaki farkı kavrayamamaları gerek biçimbilimsel gerekse sözdizimsel düzeyde hata yapmalarına neden olmaktadır. Bu tür hatalar “SDN” etiketiyle kodlanmaktadır.

33 yaşındayken iyi <err> işi </err><corr> bir iş </corr> bulmalıyım.

Her sabah duş almalıyız çünkü vücut <err> oksijeni </err><corr> oksijen </corr> istiyor.

## 5.6. Anlamsal Hatalar

Anlamsal hatalar, çoklu hata içermesi veya Türkçeye has olmayan ifade biçimlerinin kullanılması sonucu cümlenin tamamında veya büyük bir bölümünde meydana gelen anlam belirsizliklerinden kaynaklanan hatalardır. Öğrenen derleminde hata etiketleme yapılırken bozuk olmasına rağmen cümlenin anlamının bağlamdan çıkarılabildiği durumlarda “AIF” (ifade hataları) etiketi (a), cümlenin anlamının bağlamdan çıkarılamadığı veya çıkarılan anlamın doğru olup olmadığından emin olunmadığı durumlarda “AAB” (anlam belirsizliği) etiketi (b) kullanılmıştır.

a. Sağlıklı bir yaşam için iyi beslenmeliyiz çünkü <err> yeter ve doğal yemekler lazım </err><corr> doğal gıdaları yeterli miktarda tüketmek lazım </corr>...

Bana göre gelişme için en önemli <err> şeyler; birinci demokrasidir </err><corr> şeylerin başında demokrasi gelir </corr>.

b. ...<err> günümüzde her kişi hatta yalnız yapacak şey her zaman uğraşma ile yapmalıysaydı yani işleyen demir pas tutmaz yapıyor ve bundan dolayı hem kisten kisiye gelişecek hem de dünyamıza gelecek </err><corr> ? </corr>

“AACS” etiketi ise anlam bakımından birbiriyile çelişen kelimelerin cümlede bir arada kullanılmış olmasından kaynaklanan anlatım bozukluklarını ifade etmektedir.

Eğer iyi bir insansak <err> mutlaka onları hissedebiliriz </err><corr> onları hissederiz </corr>...

Burada </corr> <err> belki çok altın olmamalı </err><corr> çok altın vardır </corr>, hepsini alacağım.

## 5.7. Metin Biçimlendirme Hataları

Metin biçimlendirme hataları, yabancı öğrencilerin yazdıkları metinleri kâğıt üzerinde düzenlemeleri ve kompozisyon yazım kuralları ile ilgilidir. Başlık için ayrılan ve diğerlerine göre daha kısa bir çizgiyle gösterilen satıra başlığın yerine kompozisyonun birinci paragrafının ilk cümlesinin yazılması “MBB” (Başlık Hataları) şeklinde kodlanmıştır.

Dikkat çeken başka bir metin biçimlendirme hatası, öğrencilerden birinin her satırın başındaki kelimeyi sistematik bir şekilde büyük harfle yazmasından kaynaklanmaktadır. Bu hata, öğrencinin hem Türkçede büyük harflerin kullanıldığı yerleri bilmediğini hem de satır kavramına aşina olmadığını göstermektedir. Bu sebeple bu tür yanlış kullanımlar hem “MBS” (Satır Hatası) hem de “YBH” (Yazım/Büyük Harf Hatası) olarak etiketlenmiştir.

## 6. Tartışma ve Sonuç

Bu çalışmada, Türkçe öğrenen derleminde kullanılabilecek hata etiketlerinin nasıl geliştirildiği örnekleriyle beraber anlatılmıştır. Hata etiketleri; hata alanı (noktalama, yazım hataları vs.), hata kategorisi (hâ, çatı, eksik kullanım, yanlış kullanım vs.) ve uygun ise sözcük türü (ad, fiil, zamir vs.) olmak üzere üç boyut göz önünde bulundurularak geliştirilmiştir.

Türkçeyi yabancı dil olarak öğrenen yabancı öğrencilerin yaptığı hatalar üzerine çok sayıda çalışma (Ak Başoğul ve Can, 2014; Albayrak, 2010; Aytan ve Güney, 2015; Büyükkiz ve Hasırcı, 2013; Çerçi, Derman ve Bardakçı, 2016; Çetinkaya, 2015; Emiroğlu, 2014; Ersoy, 1997; Işloğlu, 2014; İnan, 2014; Kıvrak, 2019; Sarıca ve Od, 2015; Subaşı, 2010; Şahin, 2013; Yılmaz ve Bircan, 2015) yapılmış olmasına rağmen bu araştırmalarda öğrenen hatalarının tipolojisini yapma yoluna gidilmemiştir. Mevcut araştırmanın kapsamında geliştirilen hata tipolojisi öğrenenlerin Türkçeye, özellikle yazım kurallarıyla dilbilgisine ne ölçüde hâkim olabildiklerinin başarılı bir şekilde değerlendirilmesini sağlayabilecek, bu tipolojiye dayanan hata etiketleme sistemi ise ileride otomatik veya yarı otomatik etiketlenmiş Türkçe öğrenen derlemlerinin oluşturulmasına ivme kazandıracaktır.

Mevcut araştırmanın alanyazına katkılarının yanı sıra bazı sınırlılıkları mevcuttur. Çalışmada Türkçe öğrenenlerin dilbilimsel yeterlikleriyle ilgili olan hata etiketleri üzerinde durulmuş, söylem hataları (*discourse errors*) ise araştırmanın kapsamı dışında tutulmuştur.

Yukarıda anlatılanlarla bağlantılı olarak aşağıdaki önerileri sunmak olanaklıdır:

Türkçe öğrenenlerin yaptığı söylem hataları sınıflandırılmalı ve bu hataları kapsayacak daha geniş bir hata etiketleme sistemi geliştirilmelidir.

Ana dili konuşucularına yeterince doğal gelmeyen söylemlerin hatalı kullanımlardan ayırt edilebilmesi için kriterler geliştirilmelidir. Yukarıda bahsi geçen ifadelerin ana dili konuşucularına neden doğal gelmediği hakkında bilgi edinilmelidir.

Hangi hata türlerinin, öğrenenlerin söylemlerinin anlaşılmasını engelleyip hangilerinin engellemediği konusunda daha fazla araştırma yapılmalıdır.

---

**Hakem Değerlendirmesi:** Dış bağımsız.

**Çıkar Çatışması:** Yazar çıkar çatışması bildirmemiştir.

**Finansal Destek:** Yazar bu çalışma için finansal destek almadığını beyan etmiştir.

**Peer-review:** Externally peer-reviewed.

**Conflict of Interest:** The author has no conflict of interest to declare.

**Grant Support:** The author declared that this study has received no financial support.

---

## Kaynakça/References

- Ak Başoğul D. & Can, F.S. (2014). Yabancı dil olarak Türkçe öğrenen Balkanlı öğrencilerin yazılı anlatımda yaptıkları hatalar üzerine tespitler. *Dil ve Edebiyat Eğitimi Dergisi*, 10, 100-119.
- Albayrak, F. (2010). *Türkçe öğrenen Moğol öğrencilerin yazılı anlatım yanlışlarının dil bilgisi açısından değerlendirilmesi*. (Yüksek Lisans Tezi). Atatürk Üniversitesi Sosyal Bilimler Enstitüsü, Erzurum.
- Aytan, T. & Güney, N. (2015). Türkçeyi yabancı dil olarak öğrenen öğrencilerin yazılı anlatımlarında karşılaşılan sorunlar (Yıldız Tömer örnekleme). *International Journal of Languages' Education and Teaching*, 3(2), 275- 288.
- Büyükkiz, K.K. & Hasırcı, S. (2013). Yabancı dil olarak Türkçe öğrenen öğrencilerin yazılı anlatımlarının yanlış çözümleme yaklaşımına göre değerlendirilmesi. *Ana Dili Eğitimi Dergisi*, 1(4), 51-62.

- Çerçi, A., Derman, S. & Bardakçı, M. (2016). Yabancı dil olarak Türkçe öğrenen öğrencilerin yazılı anlatımlarına yönelik yanlış çözümlemesi. *Gaziantep University Journal of Social Sciences*, 15(2), 695-715.
- Çetinkaya, G. (2015). Yanlış çözümlemesi: yabancı dil olarak Türkçe öğrenen B2 düzeyindeki öğrencilerin yazılı metinlerine ilişkin görünüm. *International Journal of Language's Education and Teaching*, 3(1), 164-178.
- Dagneaux, E., Denness, S., Granger, S., Meunier, F., Neff, J., & Thewissen J. (Eds.) (2005). *UCL error-tagging manual, Version 1.2*. Louvain-la-Neuve: Centre for English Corpus Linguistics, Université Catholique de Louvain.
- Díaz-Negrillo, A. & Fernández-Domínguez, J. (2006). Error tagging systems for learner corpora. *Revista española de lingüística aplicada*, 19, 83-102. Erişim adresi: [https://www.researchgate.net/publication/28137922\\_Error\\_Tagging\\_Systems\\_for\\_Learner\\_Corpora](https://www.researchgate.net/publication/28137922_Error_Tagging_Systems_for_Learner_Corpora)
- Díez Bedmar, M. B. (2011). Detecting learning disorders in students' written production in the foreign language: Are learner corpora of any help? *Porta Linguarum*, 15, 35-54.
- Emiroğlu, S. (2014). Türkçe öğrenen yabancı öğrencilerin yazılı anlatımlarında Türkçenin dil bilgisi ve yazım özellikleriyle ilgili karşılaştığı zorluklar. *International Journal of Language Academy*, 2(3), 99-119.
- Ersoy, Ş. (1997). *Türkçe öğrenen yabancıların yazılı anlatım yanlışlarının dil bilgisi açısından değerlendirilmesi*. (Yüksek Lisans Tezi). Ankara Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.
- Granger S. (2008). Learner corpora. In Lüdeling, A. & Kytö, M. (Eds.) *Corpus Linguistics. An International Handbook* (pp. 259-275). Berlin & New York: Walter de Gruyter. Erişim adresi: [https://www.researchgate.net/publication/273480731\\_Learner\\_Corpora](https://www.researchgate.net/publication/273480731_Learner_Corpora)
- Granger, S. (2013). Error-tagged Learner Corpora and CALL: A Promising Synergy. *CALICO Journal*, 20(3), 465-480. <https://doi.org/10.1558/cj.v20i3.465-480>
- Hana, J., Rosen, A., Škodová, S., & Štindlová, B. (2010). Error-Tagged Learner Corpus of Czech. *Proceedings of the Fourth Linguistic Annotation Workshop*, 11-19. Erişim adresi: <https://aclanthology.org/W10-1802.pdf>
- Hunston, S. (2002). *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.
- İsiloğlu, S. (2014). Yabancı dil olarak Türkçenin öğretiminde nesne durum ekinin kullanımı ile ilgili yanlışlar ve çözüm önerileri. *Route Educational and Social Science Journal*, 1(2), 101-115.
- İnan, K. (2014). Yabancı dil olarak Türkçe öğrenen İranlıların yazılı anlatımlarının hata analizi bağlamında değerlendirilmesi. *International Periodical for the Languages, Literature and History of Turkish or Turkic*, 9(9), 619-649.
- Izumi, E., Uchimoto, K., & Isahara, H. (2005). Error Annotation for Corpus of Japanese Learner English. *Proceedings of the Sixth International Workshop on Linguistically Interpreted Corpora (LINC-2005)*, 71-80. Erişim adresi: <https://aclanthology.org/105-6009.pdf>
- Kıvrak, D. (2019). *Türkçeyi ikinci dil olarak öğrenen öğrencilerin Türkçe yeterlik algı düzeyleri ve Türkçe yazılı anlatımlarındaki yazım yanlışları*. (Yüksek Lisans Tezi). Muğla Sıtkı Koçman Üniversitesi Eğitim Bilimleri Enstitüsü, Muğla.
- López, W.C. (2009). Error analysis in a learner corpus: what are the learners' strategies? Erişim adresi: <https://www.um.es/lacell/aelinco/contenido/pdf/45.pdf>
- Meyer, Charles F. (2002). *English Corpus Linguistics*. Cambridge: Cambridge University Press.
- Nicholls, D. (2003). The Cambridge Learner Corpus error coding and analysis for lexicography and ELT. *Proceedings of the Corpus Linguistics 2003 Conference*, 16, 572-581. Erişim adresi: <http://ucl.lancs.ac.uk/publications/cl2003/papers/nicholls.pdf>

- Sarıca, N. ve Od, Ç. (2015). Yabancı dil olarak Türkçe öğretiminde tamlama algısı sorunları. *International Journal of Language Academy*, 3(1), 389-398.
- Subaşı Adalar, D. (2010). Tömer’de yabancı dil olarak Türkçe öğrenen Arap öğrencilerin kompozisyonlarında hata analizi. *Dil Dergisi*, 148, 6-16.
- Şahin, E. Y. (2013). Yabancı dil olarak Türkçe öğrenen öğrencilerin yazılı anlatımlarındaki ek yanlışları. *Tarih Okulu Dergisi*, 6 (15), 433-449.
- Yılmaz, F. ve Bircan, D. (2015). Türkçe öğretim merkezinde okuyan yabancı öğrencilerin yazılı kompozisyonlarının yanlış çözümleme yöntemine göre değerlendirilmesi. *International Journal of Language Academy*, 3(1), 113-126.

