

# Türkçe için ‘Kendi Kendine Derlem Platformu’nun (KKDP) İnternet Tabanlı Türkçe Yeterlik Sınavı (İTS) Yazma Becerisi Modülüne Eklemlenmesi ve Yazma Hatalarının Belirlenmesi Temelinde Öğrenici Derlemi Altyapısı

**Doç. Dr. Bülent ÖZKAN**  
Mersin Üniversitesi

Türkçe için Kendi Kendine Derlem Platformu (KKDP), dil arařtırmacıların arařtırma sorularına baėlı olarak biçimlendirilebilen, arařtırmacıya özel, esnek, kolay ulařılabilir, veri tabanı destekli ve arařtırma sonuçlarının saėlıklı bir şekilde raporlanabildiėi bir derlem altyapısıdır. Söz konusu platform TÜBİTAK-1005-Ulusal Yeni Fikirler ve Ürünler Arařtırma Destek Programı’nca desteklenen proje kapsamında oluşturulmuřtur.

İnternet Tabanlı Türkçe Yeterlilik Sınavı (İTS) ise adayların Türkçe becerilerini ölçmek amacıyla 2011 yılından bu yana uygulanan dil becerisi belirleme sınavıdır. Sınav adayların dil becerilerini ölçmek amacıyla sözcük ve yapı bilgisi, okuma becerisi, dinleme becerisi, yazma becerisi ve konuşma becerisi düzlemlerinde yapılmaktadır.

Bu çalışma, KKDP altyapısı üzerinden İTS’ye katılan adayların yazma becerisi bölümünde ürettikleri metinlerin KKDP’ye eklemlenmesi ve adayların yazma becerisi bölümünde yazma hatalarının sistem üzerinden nasıl analiz edilebileceėi üzerine bir tanıtım ve uygulamayı içermektedir. Bu anlamda çalışma öğrenici derlemi altyapısı örneėi olarak da düşünülebilir.

**Anahtar Kelimeler:** Kendi Kendine Derlem, Türkçe Yeterlik Sınavı, Derlem, Yazma Becerisi

## 1. Giriř

*Doėal Dil İşleme* (DDİ) olarak da adlandırılan bilgisayarlı dilbilim, *derlem* adı verilen dil örnekçelerini temel alarak, *uygulamalı dilbilimin* (applied linguistics) paralelinde, başta *sözlükbilim* olmak üzere *dilbilgisi*, *ağız*, *çeviri bilim*, *tarihsel dilbilgisi* ve *dilsel deėişim*, *dil öğrenimi* ve *öğretimi*, *anlambilim*, *kullanımbilim*, *toplum dilbilim*, *söylem çözümlemesi*, *biçembilim* ve *yazınbilim* çalışmalarında yoğunlukla kullanılmaktadır (McEnery vd., 2006 Kennedy 1998).

Bu çerçevede derlemlerin farklı arařtırma soruları için, farklı amaçlar doğrultusunda oluşturulduklarını söylemek mümkündür. Derlem uygulamalarının sözü edilen uygulamalı dilbilimin alt çalışma alanlarından biri olan *dil öğrenimi* ve *öğretimi* alanında kullanımı en belirgin şekilde öğrenici derlemleri aracılıėıyla gerçeklik bulur. Dili öğrenen bireylerin öğrenilen dilde performanslarını ve öğrenme durumlarını belirlemede öğrenici derlemleri oldukça işlevsel olarak kullanılmaktadır.

### 1.1. Amaç

Bu çalışmanın amacı, dil arařtırmacıların arařtırma sorularına baėlı olarak biçimlendirilebilen, arařtırmacıya özel, esnek, kolay ulařılabilir, veri tabanı destekli ve arařtırma sonuçlarının saėlıklı bir şekilde raporlanabildiėi bir derlem altyapısı (KKDP) üzerinden bir öğrenici derleminin nasıl oluşturulduėu ve oluşturulan derlem üzerinden yazma hatalarının nasıl belirlenebileceėini arařtırmacıların ilgisine sunmaktır.

### 1.2. Önem

Bugün dilin doėal ortamlarından yazılı ve sözlü olarak elde edilen verilerle kendine özgü kuram ve uygulamalarla ortaya koyulan derlemler, özellikle sezgisel örnekçelerden ve ikincil verilerden yalıtılmış ‘gerçek zamanlı’ dilsel verileri içermesi bakımından alana büyük katkılar sağlamaktadır (Özkan, 2010). Bu anlamda, *uygulamalı dilbilimin* yukarıda bir kısmı sıralanan alt dallarında arařtırmacıların yürüttükleri derlem temelli çalışmalar, günümüzde dil çalışmalarında deneysel sonuçlara ulařılmasına önemli katkılar sağlamaktadır. Bu çerçevede bir öğrenici derlemi üzerinden yazma hatalarının belirlenmesi alana yapacaėı katkılar açısından önemlidir.

### 1.3. Sınırlılıklar

Oluřturulan öğrenici derlemi, 2011-2015 yılları arasında İTS’ye katılan B2 düzey 49 öğrencinin İTS’nin yazma becerisi bölümünde “*Bilgisayar kullanmak çocuklara yarardan çok zarar verir? Siz bu görüşe katılıyor musunuz?*” sorusuna verdikleri cevapları içermektedir. Ayrıca yazma hataları açısından tüm metinler çalışmanın bir uygulama olması nedeniyle incelenmemiř, sadece yapılabirlik çerçevesinde bir tanıtım amaçlandığı için uygulama basamakları tanıtılmıřtır.

### 1.4. Tanımlar

#### 1.4.1. İTS nedir?

İnternet Tabanlı Türkçe Seviye Belirleme ve Yeterlik Sınavı (İTS), internet tabanlı olarak uygulanan, adayların Türkçe becerilerini ölçmek amacıyla geliştirilen Türkçe seviye belirleme ve sertifikalandırma sınavıdır. Sınav adayların *Sözcük ve Yapı Bilgisi*, *Okuma Becerisi*, *Dinleme Becerisi*, *Yazma Becerisi* ve *Konuşma Becerilerini* ölçmek amacıyla geliştirilmiş beř boyuttan olmaktadır. Ayrıca sınav yeterlilik, kur geçme ve deneme sınavı olarak 2011 yılından bu

yana uygulanmaktadır. Öte yandan, İTS "Diller için Avrupa Ortak Öneriler Çerçevesi: Öğrenim, Öğretim ve Değerlendirme" kriterleri esas alınarak hazırlanan bir sınavdır.

#### 1.4.2. KKDP nedir?

KKDP, Türkçe üzerine çalışan dil/dilbilim araştırmacıları için araştırmacıların *araştırma sorularına bağlı olarak biçimlendirilebilen, araştırmacıya özel, esnek, kolay ulaşılabilir, veri tabanı destekli ve bu çerçevede araştırma sonuçlarının sağlıklı bir şekilde raporlanabildiği* bir derlem platformudur.

Platform temelde: *Sözlük Bilim, Anlam Bilim, Sözcük Bilim, Söz Dizimi ve Biçim Bilim* olmak üzere 5 farklı modül üzerine kurulmuştur. Sözü edilen modüller, ilgili alanlarda çalışmalarını yürüten araştırmacılara derlem oluşturma, bu derlem üzerinden verileri etiketleme ve bu verileri raporlama imkanı sunan ayrı ayrı yazılım kümeleridir. Bu çalışmada ise yukarıda yer alan modüllere ek olarak İTS'de katılımcıların ürettikleri metinlerden bir öğrenci derlemi oluşturulmuş ve bu öğrenci derleminde yazma hatalarının nasıl belirlenebileceği üzerine durulmuştur.

#### 1.4.3. Öğrenci derlemi nedir?

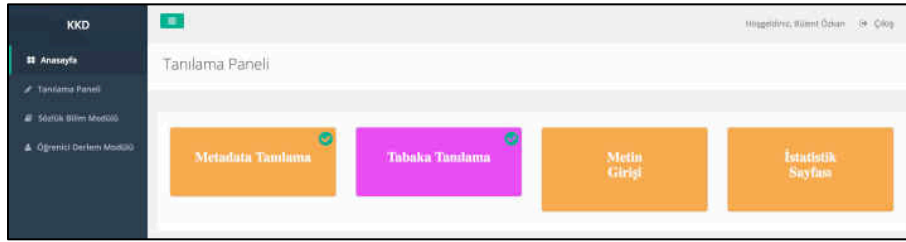
Öğrenci derlemleri (learner corpora) yabancı dil öğrencileri tarafından üretilen dilin derlem uygulamaları aracılığıyla sayısallaştırılması yoluyla oluşturulmaktadır. Bu anlamda, bir öğrenci derlemi yabancı dil öğrencilerinden elde edilen dilsel çıktılar olarak değerlendirilebilir (Baker vd., 2006). İyi düzenlenmiş bir öğrenci derlemi dillerin nasıl öğrenildiği ve dil öğrenme sürecine katkıda bulunma yollarını araştıran uygulayıcılar açısından son derece etkili bir kaynak kullanılmaktadır (Pravec, 2002).

Bu çalışmada, KKDP üzerinden örnek bir uygulamayla ortaya koyulan *Öğrenci Derlemi Modülü* yabancı dil olarak Türkçeyi öğrenenlerin yazma hatalarının belirlenmesi açısından uygulamalarla tanıtılacaktır.

## 2. Tanıtım ve Uygulama

### 2.1. KKDP öğrenci derlemi modülü tanılama paneli

KKDP üzerinde bir derlem oluşturulurken sistem üzerinde dört aşamalı bir yapılandırma söz konusudur. Bu aşamalar sırasıyla *metadata tanılama, tabaka tanılama, metin girişi* ve girilen metinlerle ilgili sayısal bilgiler sunan *istatistik sayfasıdır*. Bu aşamaların ayrıntıları aşağıda sunulmuştur (Şekil 1).



Şekil 1. Öğrenci Derlemi Modülü işlem aşamaları

#### 2.1.1. Metadata tanılama

Oluşturulacak derlem için var olan araştırma soruları çerçevesinde neleri metadata olarak belirleyeceğimiz konusu önem taşımaktadır. Bu durum derlemde metadataların raporlanma aşamasında derlemde neyi yapılandırırız onu rapor olarak alabileceğimiz temelinde düşünülmelidir (Özkan vd, 2016).

Bu anlamda öğrenci derlemi üzerinde tanılanacak metadatalar literatürle belirli ölçütlere karşılık gelmekteyse de derlem oluşturulurken belirlenebilecek temel metadatalar veri sağlanan öğrencinin: *adı soyadı, yaşı, uyruğu, cinsiyeti, bulunduğu dil beceri düzeyi, ne kadar zamandır dile maruz kaldığı, bildiği başkaca diller* gibi araştırma sorularına göre de şekillenebilen bir özellik taşır. KKDP bu anlamda esnek bir metadata belirleme olanağını araştırmacılara sunmaktadır (Şekil 2).

Şekil 2. Metadeta tanımlama arayüzü

### 2.1.2. Tabaka tanımlama

Bu panel üzerinden ayrıca **Tabaka Tanılama**, derlemde yer alan tabakaların belirlenmesine karşılık gelmektedir. Yine bu tanımlama işlemi araştırma sorularına bağlı olarak oluşturulacak bir derlemin varlığına işaret etmektedir. Bu yapılandırma da esnek bir arayüzle araştırmacının kullanımına sunulmaktadır. Örnek bir uygulama için sınırlılıklar bölümünde belirtilen biçimiyle bir tabaka tanımlama yoluna gidilmiş, soru bazlı bir tabakalama öngörülmüştür (Şekil 3).

Şekil 3. Tabaka belirleme arayüzü

### 2.1.3. Metin girişi

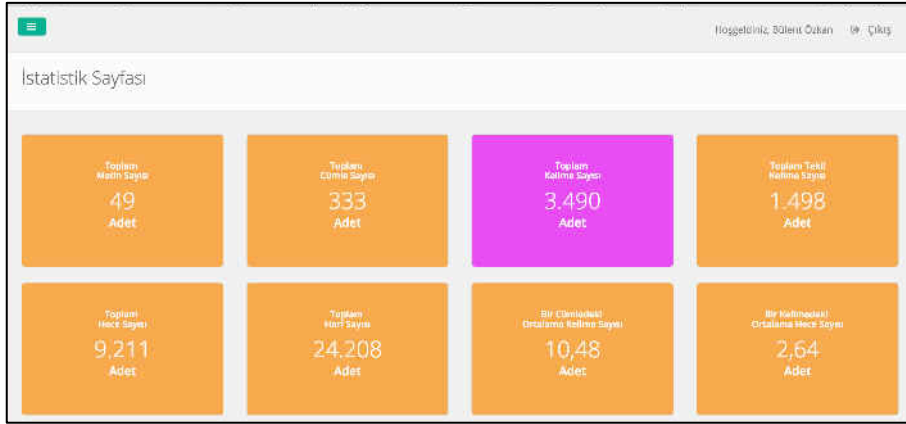
Metadeta ve tabaka tanımlama'nın ardından öğrencilerin ürettikleri metinlerin sisteme girişi bu aşamada yapılmaktadır. Bu aşamada İTS'den alınan UTF-8 biçimli metin dosyaları sisteme Metin Ekle sekmesi kullanılarak ayrı ayrı eklenebilmektedir. Metin eklemelerde belirlenen ad soyad, cinsiyet, uyruk, seviye vb. metadeta girişi de yapılabilmektedir (Şekil 4).

Şekil 4. Metin ekleme arayüzü

Metin ekleme aşamasında sistem otomatik olarak girilen metni sözbirimlerine ve cümlelere ayırmakta, metadatalar ve tabaka bilgisiyle birlikte sistem üzerinden veri tabanı kayıtlarını gerçekleştirilmektedir.

### 2.1.3. İstatistik sayfası

KKDP'ye girişi yapılan metinlerle ilgili basit/temel düzeyde sayısal veriler içeren istatistik sayfası oluşturulan derleme ilgili bilgileri kullanıcıya sunabilmektedir. Bu veriler: *toplam metin sayısı*, *toplam cümle sayısı*, *toplam kelime sayısı* gibi sayısal verilerdir (Şekil 5).



Şekil 5. İstatistik sayfası arayüzü

### 2.2. KKDP öğrenci derleme modülü metin (veri) işleme paneli

Derlemenin oluşturulmasının ardından sistem üzerinden metin (veri) işleme aşamasına geçilmektedir. Veri işleme aşaması temelde 5 alt basamaktan oluşmaktadır (Şekil 6).



Şekil 6. Metin (veri) işleme arayüzü

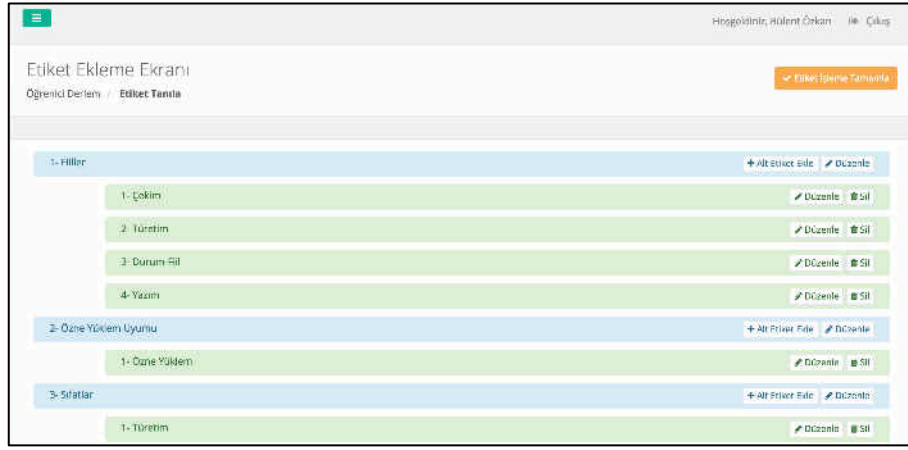
Bu aşamada sözbirimleri ve bunlara bağlı tanık cümleleri işaretlemek amacıyla **etiket tanılama** aşamasında, esnek etiketler belirlenebilmekte (Şekil 7); **sözbirimleştirme** aşamasında (Şekil 8); 'madde başılama' işlemi yapılmakta ve aynı ekranda sözbirimleştirme aşamasında hatalı yazımların gerçekte ne oldukları sorgulanarak denetlenebilmekte (Şekil 8); oluşturulan sözbirimler başka bir ekranda kontrol edilip düzenlenebilmekte (Şekil 9); **cümle düzenleme** aşamasında cümle bölümlenmelerinde hatalar mevcutsa bu hatalar giderilebilmekte (Şekil 10) ve **hata analizinde** etiket tanılama aşamasında belirlenen esnek etiketlerle işaretlemeler yapılabilmektedir (Şekil 11 ve Şekil 12). Ayrıca bu ekranda işlenen veriler raporlanabilmekte (Şekil 13 - Şekil 16) ve bu raporlamalar dışı aktarılabilir.

Metin (veri) işleme aşamasında hata analizine dahil olan başka bir durum ise **yanlış eşdizimliliklerin** belirlenmesidir. Yanlış eşdizimliliklerin belirlenmesi öğrenci derlemelerinden elde edilen başlıca bir analiz olarak literatürde yerini almıştır. Sistem üzerinden eşdizim analizleri (Şekil 17) ve bunlara bağlı olarak raporlar (Şekil 18) oluşturularak yine bu raporlar dışı aktarılabilir.

Bu aşamaların ayrıntıları şu şekildedir:

#### 2.2.1. Etiket tanılama

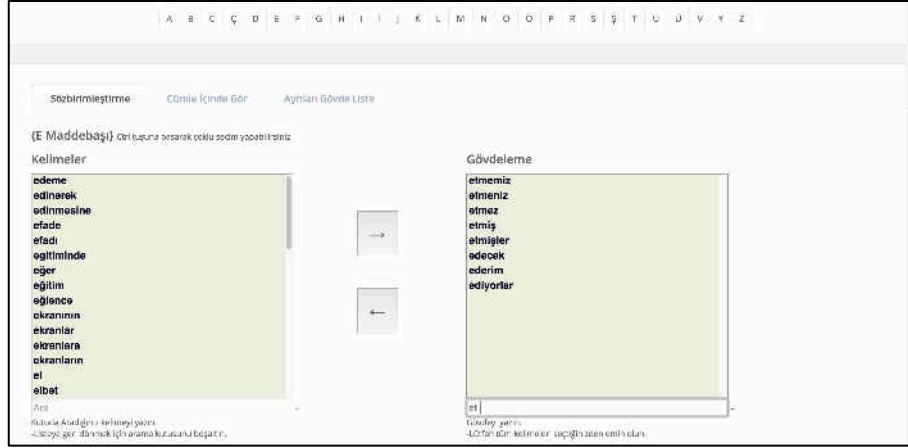
Belirlenen araştırma sorusu çerçevesinde, sözcük türleri temelinde *çekim*, *türetim*, *durum-yapı*, *yazım*; *özne-yüklem uyumu*; *tamlama* vb. nitelikler esnek etiket tanılama sistemi üzerinden belirlenmektedir. Belirlenen bu etiketler, veri işleme (hata analizi) arayüzünde sözbirimler ve buna bağlı cümle yapılarının işaretlenmesi aşamasında işlev görmektedirler (Şekil 7).



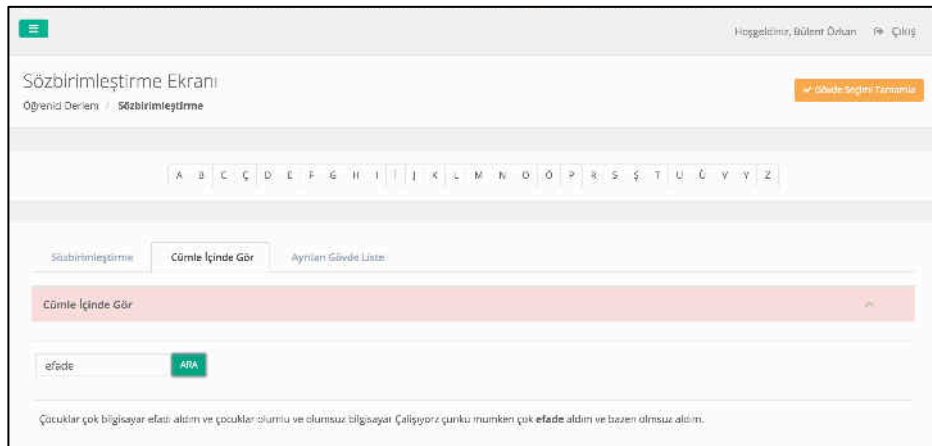
Şekil 7. Esnek etiket belirleme arayüzü

### 2.2.2. Sözbirimleştirme

Bu aşamada derlemede yer alan sözbirimler ‘madde başlama’ işlemine (sözbirimleştirme) tabi tutulmaktadır (Şekil 8). Hata analizinin dayandığı öncelikli yapı, yazım (imla) hatalarıdır. Bazı durumlarda bu yanlış yazımların sözbirimleştirilmesinde bağlam içinde görerek o sözbirimin gerçekte ne olduğuna karar vermek gerekebilir. Bu anlamda sözbirimleştirme arayüzünde bulanık yazımların da görülebileceği ‘cümle içinde gör’ aracı kullanılabilir (Şekil 9).



Şekil 8. Sözbirimleştirme arayüzü



Şekil 9. Bağlam (cümle) içinde sözcükbirim sorgulama arayüzü

Madde başlama işleminin ardından tüm çekimler bir madde başına bağlanmakta ve araştırmaya dahil edilebilmektedir (Şekil 10). Bu veri setleri düzeye göre sözcük türetim genişliğini tespit etmekte kullanılabilir.

Gövde	Kelimeler	İşlem
haha	haha, babaları, bahaların, bahaların, babası, babası	Düzenle
bağmılık	bağmılık	Düzenle
bahçe	bahçeler	Düzenle
bak-	bakırkan, bakmamasına, bakması, bakmaya, bakınıyorlar, bakıçımızda	Düzenle
bakımından	bakımında, bakımından	Düzenle
baş	baş, başına, başında, başına	Düzenle
başarsız	başarsız	Düzenle
başka	başka, başka, başka	Düzenle
başla-	başlamaları, başlar, başlayacak	Düzenle
bazen	bazarı, bazarı	Düzenle
bazi	bazı, bazı	Düzenle
beden	beden	Düzenle
belirli	belirli	Düzenle

Şekil 10. Ayrılan gövde (madde başı) görüntüleme arayüzü

### 2.2.3. Cümle düzenleme

Cümle düzenleme arayüzü sistemden kaynaklı yanlış bölümlenmeleri gidermek amacıyla oluşturulmuş bir arayüzdür. Bu arayüzde cümleler bir öncekiyle birleştirilebilmekte ya da bloklanmış cümleler bölümlenebilmektedir (Şekil 11).

Cümleler	İşlem
1- hayatınızda bilgisayar çok önemli bir alet.ve ondan vazgeçemeyiz.	Sonrakıyla Birleştir
2- Öğrenmek ve eğlenmek için kullanılabilir.	Sonrakıyla Birleştir
3- Çocuklar bilgisayarı yabancı dil öğrenilmeleri için kullanılabilir ve başka şeyler yapabilir.	Sonrakıyla Birleştir
4- Mesala ödevleri çözebilir. bazı oyuncauk oyunabilir ve ayrıca internete girbilir, arkadaşlarıyla konuşabilirler.	Sonrakıyla Birleştir
5- Ancak bilgisayarın ekranının karşısı kalırsa çok zararlı olarak ve onları gözlenne bazı sorunlar verir.	Sonrakıyla Birleştir
6- Ve onların sırtları acımaya başlayacak.	Sonrakıyla Birleştir
7- Kesarı bence: bilgisayar birim ve çocuklarımız için hem yararlı hem de zararlı iş şekilde kullanırsak yararlı olarak ve olumlu şekilde etileyecek.	Sonrakıyla Birleştir
8- Kötü şekilde kullanırsak tam tersi olarak edileyecektir	Sonrakıyla Birleştir

Şekil 11. Cümle düzenleme arayüzü

### 2.2.4. Hata analizi arayüzü

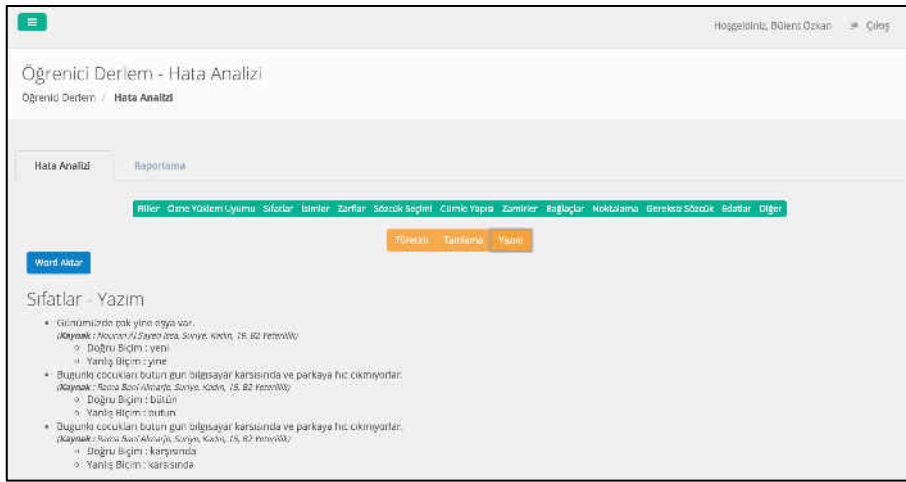
Hata analizi arayüzünde etiket tanılama aşamasında belirlenen etiketler aracılığıyla cümle ya da sözcükler doğru ve yanlış biçimleriyle sisteme kaydedilmekte (Şekil 12 ve 13) ve aynı ekranda yapılan bu kayıtlar raporlanabilmektedir (Şekil 14 - 16).

Hata Analizi	Raporlama
Tabii ki bilgisayar çocuklara çok zarar verir ve zamanın boşa kaybettirir.	İşlem: X Kaldır
Sözcük Seçimi - Sözcük   Doğru Biçim : harzahnlar, yanlış biçim : kaybetimler İsimler - Çekim   Doğru Biçim : zamanların, yanlış Biçim : zamanın	İşlem: X Kaldır
Bence hergün bilgisayar kullanılmak çocuklara yaradan çok zarar verir.	İşlem: X Kaldır
Günümüzde çok yine eyya var.	İşlem: X Kaldır
Bu günler çocuklar üç yaşında ama onunla ipad ya da bilgisayar kullanıyorlar.	İşlem: X Kaldır

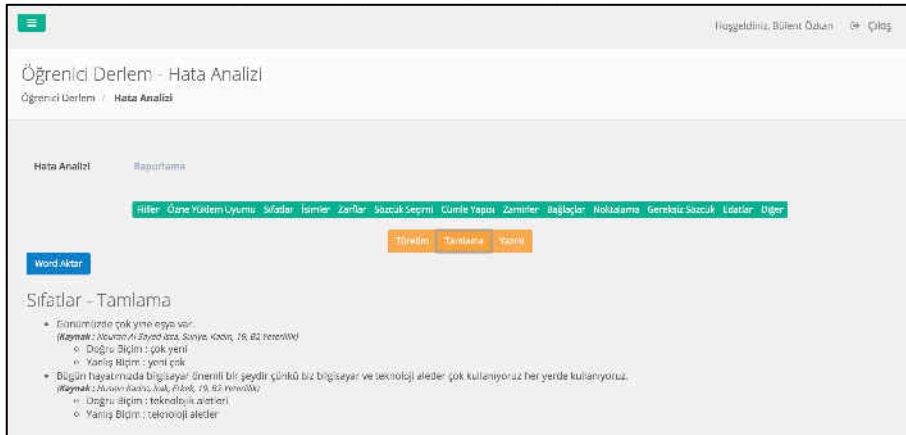
Şekil 12. Hata analizi arayüzü – I



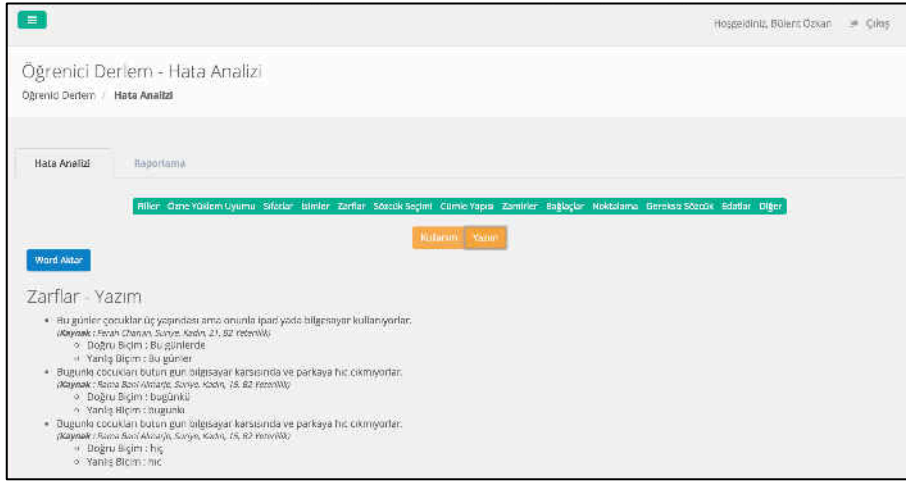
Şekil 13. Hata analizi arayüzü – II



Şekil 14. Hata analizi raporlama arayüzü (sıfat kullanımı yazım hataları)



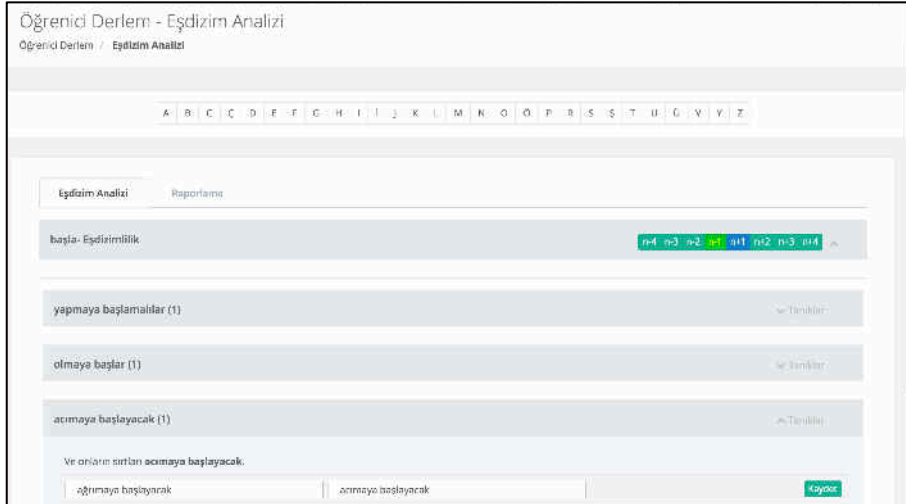
Şekil 15. Hata analizi raporlama arayüzü (sıfat kullanımı tamlama hataları)



Şekil 16. Hata analizi raporlama arayüzü (zarf kullanımı yazım hataları)

### 2.2.5. Yanlış eşdizimlik

Önce de bahsi edildiği üzere eşdizimli yapıların belirlenmesi dil öğretiminde başlı başına bir araştırma konusudur. Öğrenici Derlemi Modülü |  $n-4 \dots n+4$  | genişliğinde derlemde kayıtlı cümleleri listeleyebilmekte ve bu listelerden doğru ve yanlış eşdizimlilikler belirlenebilmektedir (Şekil 17). Ayrıca bu işlemler aynı arayüzde raporlanarak dışa aktarılabilmektedir (Şekil 18).



Şekil 17. Eşdizim analizi veri işleme arayüzü



Şekil 18. Eşdizim analizi veri işleme arayüzü

## 3. Sonuç ve tartışma

Derlem dilbilim uygulamaları bugün için *uygulamalı dilbilim* çerçevesinde karşımıza çıkan *dilbilgisi*, *ağız*, *çeviri bilim*, *tarihsel dilbilgisi* ve *dilsel değişim*, *dil öğrenimi* ve *öğretimi*, *anlambilim*, *kullanumbilim*, *toplum dilbilim*, *söylem*



çözümlemesi, biçembilim ve yazınbilim gibi dilbilimin diğer alt dallarında araştırmacılara yeni olanaklar sunmaktadır (Özkan vd., 2016).

Geliştirilen platform (KKDP), *Sözlük Bilim, Anlam Bilim, Sözcük Bilim, Söz Dizimi ve Biçim Bilim Modülü*'nün yanında, bu çalışma için geliştirilmiş olan *Öğrenici Derlemi Modülü* araştırmacılara yabancı dil olarak Türkçenin öğretimi konusunda öğretim ortamlarında aktarılan dil becerilerin geri dönütlerle beslenmesi açısından katkı sağlayamaya adaydır (Özkan vd., 2016).

Çalışma ayrıca Türkçe için geliştirilmiş olan ilk Öğrenici Derlemi altyapısı olması açısından da ayrı bir önem taşımaktadır.

Çalışmada KKDP'nin bu alan için nasıl kullanılacağına dair uygulama örneklerine yer verilmiş, dil/dilbilim çalışmalarının günümüzde olmazsa olmazı haline gelmiş bulunan bilgisayarlı dilbilim çalışmaları adına nitelikli bir uygulama örneği sunulmuştur.

### **Teşekkür**

“Türkçe için Kendi Kendine Derlem Platformu Oluşturma” başlıklı ve 114E791 proje numaralı bu çalışma, TÜBİTAK - 1005 - Ulusal Yeni Fikirler ve Ürünler Araştırma Destek Programı'na desteklenmiştir. Katkıları için TÜBİTAK'a teşekkür ederiz.

### **Kaynaklar**

- Baker, Paul, Andrew Hardie ve Tony McEnery (2006). *A Glossary of Corpus Linguistics*. Edinburgh University Press.
- Kennedy, G. (1998). *An Introduction to Corpus Linguistics*. New York: Addison Wesley Longman Limited.
- McEnery, Tony, Richard Xioa, Yukio Tono (2006), *Corpus-Based Language Studies An Advanced Resource Book*, Routledge, New York.
- Özkan B. (2010). Türkçenin Öğretiminde Sıfatların Eşdizim Sözlüğü: Yöntem ve Uygulama. e-International Journal of Educational Research. Volume 1. No 2. Autumn 2010: 51-65.
- Özkan B., Tahiroğlu B. T., Oflazer K. (2016). “Türkçe için Kendi Kendine Derlem Platformu Oluşturma” projesi sonuç raporu. Proje no: 114E791. Mersin
- Pravec, Norma A. (2002). Survey of Learner Corpora. ICAME Journal No. 26.