



## Türkçe Dilbilgisi Kurallarının Sözdizimsel Analiz Çerçevesinde İncelenmesi

Firdevs Durnagöl<sup>1</sup>, Zeynep Altan<sup>2</sup>

<sup>1\*</sup> İstanbul Beykent University, Computer Programming Department, İstanbul, Türkiye

<sup>2\*</sup> İstanbul Beykent University, Software Engineering Department, İstanbul, Türkiye

### Özet

İnsanlar kendilerini ifade etme aracı olarak doğal dillerini kullanırlar. Sözcükler öbekler halinde toplanıp bir araya gelerek ilgili dile ait tümceleri oluşturur. Öbeklerin bir araya gelmesi konuşulan dile ait belirli kurallar çerçevesinde gerçekleşir. Sondan eklemeli bir dil olan Türkçe için de biçimbilimsel, sözdizimsel, anlamsal ve sesbilimsel yapılar oluşturulmuştur. Bu çalışma kapsamında ise Türkçenin sözdizimsel analizi incelenmiştir. Çalışmada incelenen haber başlıkları içeren tümcelerin sözdizimsel analizlerine ilişkin kavramsal ifadeleri tanımlanan dilbilgisi kuralları ile ağaç yapısında betimlenmiştir. Çalışma kapsamında ifade edilen sonlu sayıda kuraldan sonsuz sayıda tümcenin üretilebilmesi mümkündür. Böylece Dil Modellemenin bir çalışma alanı olan sözdizimsel çözümleme için uygulama alanına özel teorik bir yapı çıkartılarak, otomatik çözümleme araçlarındaki belirsizlik durumlarına karşı bir çalışma yürütülmüştür.

### Makale Bilgisi

Başvuru:  
19/12/2023  
Kabul:  
06/03/2024

**Anahtar Kelimeler:** Bağlamdan Bağımsız Dilbilgisi, Kavramsal Çözümleme, Sözdizimsel Analiz.

### Examining Turkish Grammar Rules in the Context of Syntactic Analysis

### Abstract

People use their natural language as a means of expressing themselves. Words gather in groups and come together to form sentences of the relevant language. The combination of phrases takes place within the framework of certain rules of the spoken language. Morphological, syntactic, semantic and phonological structures have been created for Turkish, which is an agglutinative language. Within the scope of this study, the syntactic analysis of Turkish was examined. Conceptual expressions regarding the syntactic analysis of the sentences containing news headlines examined in the study are depicted in a tree structure with the defined grammatical rules. It is possible to produce an infinite number of sentences from a finite number of rules expressed within the scope of the study. Thus, a theoretical structure specific to the application area was developed for syntactic analysis, which is a field of study in Language modeling, and a study was carried out against the ambiguity in automatic analysis tools.

**Keywords:** Context-Free Grammar, Conceptual Analysis, Syntactic Analysis.

\* İletişim e-posta: firdevsdurnagol@beykent.edu.tr

## 1 Giriş

Dil, insanların iletişim kurmasını sağlayan bir araçtır. Dil kavramı sadece konuşulan sözcükler olarak değil, kişilerin sözcük kullanmadan kendilerini ifade etmesi durumunu da içerisinde barındırmaktadır. Kişilerin kendilerini ifade edebilmelerinde sözcüklerin yeri şüphesiz ki fazladır. Sözcükler bir araya gelerek tümceleri oluşturur ve kişilerin kendilerini ifade etmelerinde yardımcı olur. Farklı yapıdaki dillerin her birini oluşturan belirli dilbilgisi kuralları mevcuttur. Türkçenin yapısında kurallı tümcelerle birlikte diğer pek çok dilden daha fazla kuralsız tümceler vardır. Kurallar oluşturulurken dilin sahip olması gereken bütün özelliklerin göz önüne alınması oldukça önemlidir [1].

Türkçe Ural-Altay dil ailesi içerisinde bulunan sondan eklemeli bir dildir. Türkçenin sahip olduğu kurallı yapıya ait dilbilgisi kuralları geçmişten günümüze değişmemiştir. Her dilde olduğu gibi Türkçenin de temel bileşenleri sözdizimi, anlam bilgisi ve ses bilgisidir. Anlamli bir tümce kurabilmek için sözcüklerin rastgele diziliminin aksine kurallı bir yapıya ihtiyaç vardır. Kelimeler düzenli olarak bir araya gelebilmek için sözdizimi kurallarına uymalıdır. Kurallı bir yapıya göre türetilen bir dili kavramsal olarak ifade etmek mümkündür. Sözdizimsel gösterimde tümce yapısı dilin temel bileşenleri üzerine kurulur ve kelime öbeği yapısı ve ağaç diyagramı olarak iki farklı şekilde ifade edilebilir. Kurucu yapı veya kelime öbeği yapısı ve ağaç diyagramı. Chomsky [2] tüm dillerin sözdizimsel yapılarının benzer olduğunu, diller arasındaki farklılıkların, dillerin sözcüksel öğelerindeki farklılıklardan kaynaklandığını ifade etmiştir.

Çalışmada haber sitelerine özel tümcelerin sözdizimsel çözümlemesinin kavramsal olarak ifade edilmesi için, dilbilgisi teorisinden yararlanılmaktadır. Bu bağlamda çalışma alanına yönelik sözdizimi kuralları türetilmekte ve gazete manşetlerinden bu kurallara uygun örnekler çözümlenmektedir. Tanımlanan sonlu sayıdaki dilbilgisi kuralı sonsuz sayıda tümcenin sözdizimsel çözümlemesinin yapılması mümkün olacaktır [3].

Sonraki bölümde Türkçe için yapılan Doğal Dil İşleme uygulamalarına ilişkin genel bir literatür araştırması yapıldıktan sonra, Bölüm 3'de Doğal Dil İşleme uygulamaları sınıflandırılmakta ve örnekleri ile açıklanmaktadır. Uygulama alanına ait örnek tümcelerin dilbilgisi kurallarına göre türetilen

sözdizimsel ağaçları Bölüm 4 'de çıkarılmıştır. Bölüm 5 günümüz uygulamalarında doğal dil işlemede sözdizimsel analizinin önemini vurgulayan sonuç bölümüdür.

## 2 Literatür araştırması

Oflazer [4] yapmış olduğu "Türkçe ve Doğal Dil İşleme" adlı çalışmasında, Türkçe'nin sahip olduğu dil yapılarına ait özellikleri özetlemiştir. Bu bağlamda tümcenin öğelerine ayrılması biçimbirim ve sözdiziminden oluşur. Böylece Türkçe'nin sözdizimsel özelliklerinin sınırları çizilmiştir. Bu sınırlar belirlenirken Türkçe'nin sahip olduğu eklemeli yapısı nedeni ile başka dillere çeviri sırasında yaşanabilecek zorluklardan bahsedilmektedir. Çalışmada bu problemlere çözüm getirilerek Türkçe-İngilizce çeviri örnekleri verilmektedir. Oflazer [5] bir başka çalışmasında, Türkçenin morfolojik analizini gerçekleştirmiştir. Bir dilin verimli olarak çözümlenebilmesi için, sözdizimsel analizden önce morfolojik (biçimbirimsel) analizine gerek vardır. Oflazer Türkçenin biçimbirimsel analizini iki düzeyli bir yapıda tasarlamış, dil çözümleyicisi olarak Xerox sonlu durumlu makinasını kullanmıştır. İki düzeyli morfolojik analizde, birinci seviye fonetik düzeyi betimler ve dile ilişkin birimleri ve ses yapısını inceler, ikinci seviyede ise sözcüklerin kök ve çekim ekleri incelenir [6].

"The Role of Morphological Analysis in Natural Language Processing" isimli çalışmada Türkçe eylemlerin biçimbirimsel işleyişi iki seviyeli morfoloji yapısında hem üretici hem de çözümleyici kurallar ile tanımlanmıştır [7]. Çalışmada Türkçe sözcüklerdeki ünsüz ve ünlü harflerin nitelikleri görsel ve kavramsal olarak iki düzeyli morfoloji ile tanımlanmaktadır. Böylece Türkçe eylemlerin oluşturulması ve çözümlenmesi biçimbirimsel olarak gerçekleştirilmiştir. Diğer bir çalışmada ise Türkçe tümcelerin sözdizimsel çözümlemelerindeki belirsizliği gidermek üzere bir etmenlenmiş dilbilgisi algoritması tanımlanmıştır [8]. Çalışmada örnek bir tümcenin biçimbilimsel analizi yapılarak eylemin kökü çıkartılmış ve tümce bağlamdan bağımsız dilbilgisi kurallarına göre çözümlenmiştir. Daha sonra aynı tümceye etmenlenmiş dilbilgisi kuralları uygulanarak tümcedeki belirsizliği gidermek üzere ağaç yapısı çıkartılmıştır.

Dönmez ve Adalı [9] yapmış oldukları "Türkçe Tümce Çözümlemede Vektör Yaklaşımı" isimli çalışmalarında Türkçenin anlamsal yapısını incelemektedir. Ayrıca tümce kavramının öbek

yapısına değinilerek, sözdizimsel anlamda bir tümcenin hangi öbek yapılarından oluştuğunun sınırları ifade edilmiştir. Bu, bir tümcenin anlam ifade edebilmesi için sözdizimsel yapısı ile birlikte anlamsal çözümlenmesinin de birlikte incelenmesinin gerektiğidir.

Bilgin vd. [10] yapmış oldukları çalışmada, "Balkanet" projesi kapsamında yaklaşık 15 bin adet Türkçe kelimenin eş anlamlılarını çıkaran bir uygulama geliştirmişlerdir. Projede "Wordnet" uygulamasının Türkçe için bir prototipi geliştirilmiştir. Wordnet uygulaması iki yüzden fazla anlamsal ilişkinin bulunduğu kelimeler barındırmaktadır [11]. Uygulama, belirli gruplamalar yaparak anlamsal olan yakın kelimeleri eş anlamlı kabul eder.

Türkçedeki yazım yanlışlarının Doğal Dil İşleme ile giderilmesi amacıyla yapılan bir tez çalışmasında tümce yapıları incelenmiştir [12]. Çalışma kapsamında tümce içerisindeki sözcüklerin eklerinin bulunması ve eklerinin ayrıştırılması işlemleri gerçekleştirilmiştir. Böylece tümcenin anlam bütünlüğünün sağlanması hedeflenmiştir. Araştırma sonucunda, Türkçe sözlük içerisinde yapılan kelime hatalarına ilişkin önerme yapılmıştır.

"Building a Lexical Functional Grammar for Turkish" isimli doktora tezinde "ParGram" isimli bir proje geliştirilmektedir [13]. Geliştirilen proje Türkçenin diğer dillerle karşılaştırmalı olarak çözümlenmelerini derinlemesine analiz etmektedir. Proje, Türkçe bir metindeki her kelimenin kökenine inerek metin içerisindeki bütün dilbilgisi kurallarına erişebilmeyi sağlar; ayrıca eklenti olarak "LingBrowser" içerisine eklenir. LingBrowser, herhangi bir metindeki kelimelerin morfolojik, telaffuz ve anlamsal özellikleri ile ilgili bilgilendirme yapan otomatik bir araçtır.

Şahin ve Amasyalı [14] çalışmalarında metin içerisinden bilgi çıkarımı işlemini şablonlarla gerçekleştirmişlerdir. Geliştirilen uygulama Türkçenin biçimbirimsel yapısını sınıflandırmakta ve yeni bir yapı geldiğinde bu sınıflandırmaya eklemektedir. Sınıflandırma yapılırken biçimbirim ile anlam bilgisi arasındaki ilişki göz önüne alınmıştır. Yapılan çalışmada iki çeşit veri seti mevcuttur. Bunlar morfolojik analizlerin yapıldığı ve yapılmadığı veri setleridir. Çalışma sonunda şablonlar yöntemi kullanılarak morfolojik analizlerin yapıldığı veri setlerinde daha iyi şablonlar elde edilmiştir.

"Türkçe Derlemler için Sözdizimsel Görselleştirme ve Sorgulama Aracı" isimli çalışmada Türkçe derlemler ve Türkçe bir derlemin etiketlenmesi konuları araştırılmaktadır [15]. Sorgulamalar Türkçe derlemlerden yararlanarak gerçekleştirilmiş ve derlemlerdeki tümcelerin sözdizimi ağaçlarını sözcük türleriyle birlikte görselleştiren bir yazılım aracı geliştirilmiştir.

Türkçe için yapılan çalışmalar incelendiğinde her çalışmanın bir öncekini geliştirdiği ve Türkçenin dilbilimsel yapısının daha açık ifade edilebildiği görülmektedir. Yapılan farklı çalışmalarda farklı yaklaşımlar Türkçe dilindeki sözdizimsel analiz sınırlarını farklı açılardan ele alarak tümcelerin ifade edilme şekillerini sınırlandırmıştır.

### 3 Doğal Dil İşleme Uygulamalarının Sınıflandırması

Doğal Dil İşleme kavramı bilgisayarın dili anlaması ve ortama aktarması açısından önemlidir ve günümüzde birçok aktif uygulaması bulunmaktadır. Doğal Dil İşleme uygulamaları dili anlama, konuşma ve sözdizimi problemlerini çözerek araştırmacılara yardımcı olmakta ve dilin kullanıldığı çalışma alanlarında pek çok fayda sağlamaktadır. Örneğin, bir metin yığını içerisinden anlamlı bilginin çıkartılması ve anlamlı bilgi üzerinden çıkarım yapılması hem çalışma zamanını düşürmekte hem de uygulamanın doğruluğunu ve kalitesini yükseltmektedir. Doğal Dil İşleme uygulamaları genel olarak aşağıdaki şekilde gruplandırılabilir.

#### 3.1 Metin sınıflandırma

Doğal Dil İşleme'de metin sınıflandırma, temel gruplandırma olarak alınabilir. Gündelik hayatta veriler dağınık bir şekilde bulunurlar. Metin içerisinden anlamlı bir bilginin çıkarılabilmesi için metnin amacına uygun olarak sınıflandırılmış, yani etiketlenmiş olması gerekir. Bu sebeple metin sınıflandırma Doğal Dil İşleme'nin ilk adımudur. Haber sitesi içerisinde başlığı verilmeyen bir haber geldiğinde bu haberin hangi başlığa ait olduğunun tahmin edilmesi örnek olarak verilebilir. Bu örneğin metin sınıflandırması belirli etiketlerle, yani denetimli öğrenme ile yapılır. Denetimli öğrenme algoritmalarından yararlanılarak metin içerisinden anlamlı bilgiler çıkarılır. Bu işlem yapılırken kelimelerin tek tek köklerine inilir ve farklı algoritmalarla sınıflandırma işlemi tamamlanır. Bir kelimenin hemen hemen tüm dillerde birden fazla anlamı bulunmaktadır. Bazı durumlarda kelimenin sözcük yapısı olumlu iken tümceye olumsuz anlam katabilmektedir.

Metin sınıflandırma yapılırken duygu analizi ile birlikte kullanılır. Twitter üzerinden Türkçe içeriklerin incelendiği örnek, bir çalışma kapsamında metin sınıflandırma ve derin öğrenme algoritmalarından yararlanmıştır [16]. Çalışmanın amacı sosyal medya aracılığı ile yapılan yorumlar üzerinden BERT (Bidirectional Encoder Representations from Transformers), CNN (Convolutional Neural Network), LSTM (Long Short-Term Memory) sinirsel öğrenme algoritmalarını karşılaştırarak en iyi başarı gösteren sınıflandırıcının tespiti. Mevcut veri seti için BERT algoritmasının daha iyi başarı oranına sahip bir sınıflandırıcı olduğu sonucu elde edilmiştir.

### 3.1.1 Duygu analizi

Metin sınıflandırmaya yardımcı olan özelliklerden biri duygu analizidir. Duygu analizi sadece kelimelerin kök anlamlarını değil, kelimelerin sahip olduğu mecaz anlamları, olumlu-olumsuz anlamları, samimiyet, korku, öfke gibi duyguları incelemektedir. Duygu analizi, sınıflandırma yöntemlerinde sıkça kullanılmaktadır. Analizler, genellikle sinirsel öğrenme kullanılarak gerçekleştirilmektedir [16].

Bir metin içerisinde duygu analizi yapılırken genelde üç duygu incelenir; bunlar olumlu, olumsuz ve tarafsız olarak sınıflandırılır. Bu yolla çözümlemede metin içerisinde kelimelerin tek tek ve birlikte anlamları göz önüne alınır (n-gram yöntemi). Böylece kelimelerin tek başına ya da birlikte kullanıldıklarında, anlamlarının anlaşılması kolaylaşır. Türkçe metinlerde duygu analizi yapılırken sıklıkla Zemberek [17] kütüphanesinden yararlanılmaktadır. Zemberek kütüphanesi Java dilinde geliştirilmiş açık kaynak kodlu bir Türkçe dil kütüphanesidir. Bu kütüphane, kelime kökü çıkarma, dilbilgisi analizi, tümce ayrıştırma, yazım denetimi, kelime türetme gibi çeşitli işlemler gerçekleştirmektedir. Duygu analizi yapılırken de Zemberek gibi hazır kütüphanelerden yararlanılarak daha duyarlı analizler yapılabilmektedir [18].

Örneğin, "Web ortamındaki yazılı sınavların Doğal Dil İşleme yöntemleri ile değerlendirilmesi" isimli çalışmada, metnin sınıflandırılması için derin öğrenme yöntemlerinden yararlanılmıştır [19]. WordNet ve Zemberek kütüphanelerinden yararlanılarak Türk dili içerisindeki anlamsal ilişkiler oluşturulmuştur. Geliştirilen sistemin başarı oranı %92 olarak tespit edilmiştir. Bu

başlamda sınav değerlendirmesi için geliştirilen sistemin kullanılması yeterli görülmüştür.

### 3.1.2 Yanlılık ve sahte haber tespiti

Günümüzde en büyük bilgi güvenilirliği sorunlarından bir tanesi içeriklerin sahte ya da yanlış olmasıdır. Haberlerin sahte ya da yanlış yazılması okuyucuları yanlış yönlendirmekte ve bilgi kirliliği oluşturmaktadır. Bunların tespit edilmesi ve doğru bilginin okuyuculara ulaştırılması önemlidir. Bu da ancak yanlılık ve sahte haber tespiti ile mümkündür. Sahte haber tespitleri yapılırken metin sınıflandırma yönteminden yararlanılmaktadır. Sahte haberlerin tespiti için yapılan çalışmada bir sınıflandırma yapılmakta ve Twitter haberleri üzerinden sahte haberler tespit edilmektedir [20].

Örneğin, sahte haberlerin güvenilirliğinin artırılması için birçok yapay zekâ uygulamasından yararlanılmaktadır. Bunlardan bir tanesi Deepfake teknolojisidir. Deepfake teknolojisi ile video içerisine ses ve görüntü eklenmektedir. Bu durum teknik olarak incelenmediği sürece anlaşılmamaktadır. Kişilerin yüzleri değiştirilmekte ve yanlış ya da sahte bir haber inandırıcılığı yüksek bir şekilde yayılmaktadır. Örneğin, Ziraat Bankasının reklam filminde, Deepfake teknolojisi kullanılarak vefat etmiş olan oyuncu Kemal Sunal'ı oynatmıştır.

Bir başka örnek ise, 2016 yılında ABD'de yapılan seçim kampanyasında yapay zekâdan yararlanılarak seçmenlerin yönlendirilmesidir. Kullanılan "mikro hedefleme stratejisi" ile oy verenler gruplara ayrılarak hedef gruplar belirlenmiş ve bu gruplara göre ayrı ayrı seçim vaatleri yapılmıştır [21].

### 3.2 Makine çevirisi

Makine çevirisi, bilgisayar aracılığı ile bir dilin başka bir dile çevrilmesidir. Makine çevirisi yapılırken Doğal Dil İşleme yöntemlerinden de yararlanılabilir. Çeviri için kural tabanlı çeviri, istatistiksel makine çevirisi ve sinirsel makine çevirisi olarak sınıflandırılan üç temel yöntemden biri seçilir. Makine çevirisinin tarihsel gelişimi araştırıldığında, 1990'lı yılların sonlarında kural tabanlı makine çevirisi yapan uygulamalara rastlanmaktadır. Systran, Lionbridge uygulamaları bu yöntemi kullanan uygulamalar arasında yer almaktadır [22]. 2006 yılına gelindiğinde, veri tabanı geliştirilerek kural tabanlı çeviriden daha iyi bir sonuç veren istatistiksel makine çevirisi kullanılmıştır. Google, Microsoft uygulamaları

belirtilen yıllarda istatistiksel çeviri yapan yöntemleri uygulayan çeviri programlarını kullanırlar. Teknolojinin de ilerlemesi ile birlikte çeviri yöntemleri de gelişmiş, 2016 yılında sinirsel makine çevirisi kullanılmaya başlanmıştır. Günümüz uygulamalarında da yaygın kullanılan yöntem sinirsel makine çevirisi yapmaktır [23]. DeepL, Google çeviri programları günümüzde kullanılan sinirsel makine çevirisi uygulamaları arasında yer almaktadır.

Babelfish uygulaması ilk olarak kural tabanlı geliştirilmiştir. 2008-2012 yıllarında Microsoft ile yapılan iş birliğiyle çeviri yöntemleri geliştirilerek istatistiksel çeviri yöntemi kullanmış ve uygulamanın adı Bing Çeviri olmuştur. 2016 yılı sonrasında sinirsel makine çevirisi kullanılarak çevirilerin daha doğru yapılması sağlanmış ve Microsoft Çeviri adını almıştır [24].

Google Çeviri uygulamasında istatistiksel tabanlı çeviri yapılmıştır. Daha sonra uygulama sinirsel makine çevirisi olarak güncellenmiştir [23]. Çalışmada istatistiksel ve sinirsel makine çevirilerinin karşılaştırmaları biçimbirimsel, sözcükbirimsel sözdizimsel ve anlamsal analiz bakış açılarından gerçekleştirilmiştir. Sonuç olarak, sinirsel makine çevirisindeki gelişim ve çevirinin yapısal doğruluğu ortaya konulmuştur.

Başka uygulama olan DeepL, günümüzde kullanılan ve Google Çeviri uygulamasına benzeyen bir çeviri uygulamasıdır. Google Çeviri programı gibi sinirsel makine çevirisi yönteminden yararlanarak çeviriler yapmaktadır. 2023 yılında yapılan çalışmada [25] Google Çeviri ve DeepL uygulaması performansı karşılaştırılmıştır. Sonuç olarak anlatımcı ve işlevsel metin çevirisinde DeepL uygulamasının Google Çeviri uygulamasına göre daha iyi sonuçlar verdiği çıkarılmıştır.

Bir başka çalışmada dokümanların boyutlarına göre makine çevirisi yapan uygulamalar arasından ChatGPT incelenmiştir [26]. Günümüzde yaygın kullanılan ChatGPT'nin karşılaştırmalı çeviri analizlerinde ortalama bir başarı gösterdiği ifade edilerek, söylem bilgisinin açıklanması konusunda daha iyi olduğu vurgulanmaktadır. Çalışma, insan çevirisinden ziyade makine çevirisinin daha doğru ve güçlü olduğunu ortaya koymaktadır.

Çeviri yöntemleri ile ilgili ayrıntılı bilgiler aşağıda verilmektedir.

### **3.2.1 Kural tabanlı çeviri**

Otomatik çeviri programlarından çok önceleri kullanılan bir çeviri yöntemidir. Dilin sahip olduğu kurallar göz önüne alınarak ve bu dilin kuralları ayrıştırılarak makine diline çeviriler yapılırdı. Bu yöntem tercih edildiğinde yapısal ve dil bilgisi özellikleri ortak olan diller arasında kullanılırsa daha iyi sonuç elde edilmektedir. Fakat her dilin kendi içerisinde sahip olduğu farklı karmaşık yapıları ve çift anlamlı kelimeleri bulunmaktadır. Bu sebeple uygulamalarda K kural tabanlı çeviri yönteminin birebir kullanımı fazla tercih edilmemektedir [27].

### **3.2.2 İstatistiksel makine çevirisi**

İstatistiksel dil modelleri kendinden önceki tümcelerden yola çıkarak kendi durumunu tahmin etme üzerine dayanmaktadır. Daha sonra geliştirilen yöntemler kelimeleri de işin içerisine katmaktadır. Bu sınıflandırmadaki modeller, aynı aileden gelen diller arasındaki çeviriyi yaparken daha başarılı olmaktadır. Bu başarı, aynı dil ailesine sahip dillerin sözdizimsel yapılarından kaynaklanmaktadır [28]. Fakat dil ailesi genişledikçe, yapısal farklılıklar ortaya çıktığından çevirilerden de yeterince doğru sonuçlar alınamamaktadır. Bu gruptaki çeviri yönteminde büyük verilerle çalışılarak veriler kümelerle ayrılır ve kümeler üzerinde istatistiksel yöntemler kullanılarak çeviri yapılır. Kaynak dilin hedef dile çevrilmesi esnasında çeviri modelleri eğitilerek çeviri geçişlerinin olasılık değerleri belirlenir. Büyük veriden elde edilen olasılık değerleri üzerinden çeviriler tamamlanır. Ancak sonuçlar tam anlamıyla dilbilgisi kurallarını ve tümce yapısını anlayamaz [29]. Bu durumun ortadan kaldırılması için, sinirsel makine çevirisi yöntemine başvurulmaktadır.

### **3.2.3 Sinirsel makine çevirisi**

Makine çevirisi yapılırken istatistiksel modellerin eksiklerini kapatacak nitelikteki çeviri yöntemidir. Bu yöntemde tümceler tek bir bağlam çerçevesinde ele alınır. Tümcelerden kelimelere inilerek tüm bağlam eş zamanlı çevrilir. Bu durum istatistiksel makine çevirisinde yaşanan sorunu çözme niteliğindedir. Tümce içerisindeki kelimeler diğer kelimelerde eşleştirilerek bir kombinasyon oluşturulur. Oluşturulan kombinasyon sonucunda sinir ağı modelleri ortaya çıkar ve çeviriler daha doğru şekilde yapılır. Günümüzde aktif olarak birçok makine çeviri programı kullanılmaktadır. Bu programlar genel olarak sinirsel makine çeviri

yöntemi tabanlıdır. Makine çeviri programları bir çevirmenin uzun bir sürede yaptığı işi çok kısa sürede tamamlamaktadır. Bazı uygulamalar ücretsiz olarak sunulduğundan daha çok tercih edilmektedir. Microsoft Bing, Google Çeviri bu uygulamalara örnek olarak verilebilir [30].

### 3.3 Öngörmeli metin modelleme

Öngörmeli metin modelleme sistemlerinde tümceenin tamamı yazılmadan ilk kelimedenden itibaren kelime tahmini yapılarak olası kelimelerle tümce tamamlanır. Birçok uygulama tarafından sunulan bu modellemeye örnek olarak, Microsoft'un kullandığı Swiftkey klavye uygulaması [31] ve Fleksy klavye uygulaması verilebilir [32]. Özellikle metin tabanlı uygulamalarda sıkça kullanılarak, kullanıcılara olası durumlar gösterilip seçenekler sunulmaktadır.

Örneğin, Google'ın geliştirdiği web tabanlı yazım uygulaması Google Docs, temel dilbilgisi özelliklerine ve bir yazım denetleyicisine sahiptir. Yazı yazarken hataları düzelterken otomatik düzeltme seçeneği vardır. Ayrıca, kelimenin ne olacağını tahmin eden metin tahmini özelliğine sahiptir. Tüm bu özelliklerine rağmen, Google Docs her şeyi doğru tahmin edemez. Uygulamada ne kadar çok yazı varsa sistemin önerileri o kadar iyileşir; böylece hem daha fazla metin yazılmasına olanak sağlanır, hem de neyi nasıl kullanmanın daha doğal görüneceğine ilişkin iyi bir ipucu verilmiş olur.

Bir diğer örnek ise arama motorlarıdır. Günümüzde sıklıkla kullanılan popüler arama motoru Google, tahminli tavsiye sistemine sahiptir. Google arama çubuğuna bir kelime yazıldığında, bu kelime ile ilişki diğer kelimeler belirir ve istenilen kelime seçilerek arama sonucuna daha isabetli erişilebilir.

### 3.4 Soru cevap sistemleri

Soru cevap sistemleri belirli bir veri tabanı içerisinde kullanıcıların sorgularını yaptığı ve bu sorguları yaparken günlük kelimelerden yararlanıp sınırlama olmaksızın sorgulamaları gerçekleştirdikleri bir sistemdir. Bu sistemde anahtar kelimeler belirlenir ve kelimelerin sıklıkları veri tabanı içerisindeki kelimelerle karşılaştırılır. Sorgulama işlemi esnasında kelime sözcük bazlı arama yapılır; kelimedenden önce ve sonra gelebilecek olan kelimeler tahmin edilerek bağlam oluşturulur. Teknolojinin hızla gelişmesiyle birlikte soru cevap sistemleri de gelişmiştir. Sadece veri tabanında bulunan hazır soruların cevapları yerine, günümüzde kullanılan sinirsel ağ temelli akıllı uygulamalar müşterilerin bütün sorularına cevap

verebilmektedir ve müşterilerin işlemlerini arada bir insan olmadan hızlı bir şekilde gerçekleştirmektedir [33]. Örneğin, Garanti bankasının kullanmış olduğu UĞİ akıllı asistan uygulaması banka müşterilerinin sordukları soruları, arada bir aracı olmaksızın, hızlı bir şekilde cevaplamaktadır. Sinirsel ağ temelli bu uygulama sayesinde hem maliyetten hem de zamandan tasarruf edilmektedir [34]. Bu tarz uygulamaların kullanım alanı geniştir. Banka sektörünün yanı sıra tıp sektörü, seyahat sektörü başta olmak üzere birçok sektörde aktif olarak kullanılmaktadır. Kişiler soru cevap sistemlerini hem yazılı olarak hem de konuşarak kullanabilirler. Bu uygulamalara diğer örnekler, sorularla kişinin hastalığını teşhis eden yapay zekâ uygulamaları, otellerin kullandığı akıllı seyahat asistanları, akıllı çağrı merkezi elemanları verilebilir.

Arama motorları, soru cevap sistemleri içerisine klasik bir örnek olarak girer. Kullanıcılar sorguladıkları sorulara cevaplarını arama motorları sayesinde bir veri tabanı içerisinden alırlar. Örneğin, ASKMSR soru cevap sisteminde, sorgular yeniden oluşturularak Google veri tabanı içerisinden cevaplanır. Cevaplamalar yapılırken kelime bazlı frekans ölçümlenmeleri yapılmakta ve tekli, ikili, n'li sözcüklerin frekansları üzerinden kelimelerin yan yana gelme olasılıkları hesaplanarak karar verilmektedir [35].

### 3.5 Metin özetleme

Mevcut bilginin fazlalığı ve giderek artması istenilen bilgiye ulaşmayı geciktirmekte ve zaman zaman imkânsız hale getirmektedir. Bilgi yığınının içerisinden istenilen temel bilgiye ulaşmak günümüzün sorunlarından bir tanesidir. Bu sorunun çözümü bilgiye doğru erişim ile sağlanmaktadır. Bilgiye erişmek için metin özetleme yöntemi kullanılabilir.

Metin özetleme, mevcut bilgi içerisindeki önemli kısımların araştırmacıya ulaşmasını sağlayarak, kişinin genel bağlamdan bilgi sahibi olmasını sağlamaktadır. Özetleme yöntemleri Sorgu Tabanlı Özetleme, tek ve çok kaynak tabanlı Kaynak Özetleme olarak farklı bakış açılarından sınıflandırılabilir. Sorgu tabanlı özetleme, yöneltilen sorgu çerçevesinde yapılan özetlemelerdir. Örneğin, Textrazor [36] uygulaması gelişmiş bir metin özetleme uygulamasıdır. Sahip olduğu özelleştirme sayesinde metni özetlerken istenilen bilgi türünde bir iki kelime girilmesi durumunda kavramsal olarak kelimelere anlamsal

olarak benzeyen model oluşturulmasını sağlar. Tek veya çok kaynaklı özetleme ise, birden fazla kaynağı referans alarak yapılan özetlemelerdir. Örneğin, SummarizeBot [37] uygulaması metin özetleme yaparken tek ya da birden fazla kaynak kullanarak çalışır.

Metin özetleme çalışmalarının ilk dönemlerinde istatistiksel yöntemlerden yararlanılmaktaydı. Kullanılan yöntemler zamanla maliyet ve başarı açısından yetersiz kaldı. Günümüzde sinirsel ağ tabanlı metin özetleme yöntemleri kullanılmaktadır. Örneğin, Pegasus uygulaması Google tarafından üretilen metin özetleme uygulamasıdır. Mevcut uygulama Google'ın daha önceden geliştirmiş olduğu BERT modeli tabanlı çalışmaktadır. BERT modeli, denetimsiz öğrenme ile geliştirilen bir sinir ağı modelidir ve metin özetleme için özelleştirilmiştir. BERT modelinin kullanmış olduğu denetimsiz öğrenme yöntemi sayesinde Pegasus uygulaması sadece metinleri özetlemeye odaklanmaz, metni anlamaya çalışır. Uygulama kendi içerisinde metin üretim modeli de kullanarak önemli vurgulanması gereken kısımları daha iyi analiz eder ve bu şekilde daha iyi özetleme işlemi gerçekleştirir [38].

#### 4 Haber başlıklarının sözdizimsel analizleri

Tümce içerisindeki sözcüklerin sözdizimsel fonksiyonları tümcenin sahip olduğu işleve göre değişmektedir. Bir tümce içerisinde bir veya daha fazla isim, sıfat, fiil, zarf ve bağlaç bulunabilir. Bu etmenler birleşerek tümcenin tamamını oluştururlar. Bu bölümdeki sözdizimsel analizlere ait Türkçe dilbilgisi kuralları, ilk olarak ad öbeği ve eylem öbeğine ait kurallar olarak tanımlanmış; daha sonra bunlar Türkçe dilbilgisi kuralları çerçevesinde bir araya getirilerek tümceler oluşturulmuştur. Tanımlanan Türkçe dilbilgisi kuralları ve sınırlamaları aşağıda verilmektedir.

##### 4.1 Sözdizimsel kuralların sınırlarının oluşturulması

Herhangi bir Türkçe anlamlı tümcenin oluşturulabilmesi için, belirli dilbilgisi kurallarına uyulması gerekir. Bu kurallar tümcenin sözdizimsel yapısı ile birlikte anlam bilgisini de içerir. Bu nedenle Doğal Dil İşlemede tümcelerın sözdizimsel analizi ön plana çıkmaktadır. Sözdizimsel analiz için sonlu sayıda dilbilgisi kuralına ihtiyaç vardır. Bu kuralların sınırlarının belirlenmesi, öbek yapısının belirlenmesi ile gerçekleşir. Türkçe 'de kurallı tümceler, ad öbeği ile eylem öbeğinin birleşmesi ile oluşmaktadır.

Ad öbeği yapısal olarak ya tek başına ifade edilebilir ya da kendinden önce veya sonra gelen adlarla belirli kümeler oluşturabilir. Türkçe'de niteleyici sözcükler ad öbeği betimlemesi içerisinde değerlendirilirler. Eylem öbeğinin yapısal tanımlaması ise, ad öbeği sınırları içerisinde girmeyen eylem bildiren sözcüklerin içerilmesidir.

Türkçe tümce yapısını oluşturan dilbilgisi kuralları ve bunların farklı kombinasyonları aşağıda açıklanmaktadır:

i) Tümce, öbek halinde bir yapı olarak Ad Öbeği (AÖ) ve Eylem Öbeği (EÖ)'nün art arda eklenmesinden oluşur. C ile betimlenerek

$$C \rightarrow AÖ EÖ$$

temel dilbilgisi kuralını oluşturur. Tümcesi giriş tümcesine ait olmayan terminal (giriş) olmayan sembollerin sonlu kümesi N ve terminal (giriş) sembollerinin sonlu kümesi  $\Sigma$  ile aşağıdaki şekilde elde edilir.

$$\Sigma \rightarrow \{\text{Eylem, Sıfat, Zarf, Ad, Eki, Bağlayıcı, Ünlem, Rakam}\}$$

$$N \rightarrow \{C, AÖ, EÖ, EKLER, BİREylem, CDU, Sayı\}$$

$\Sigma$  kümesinin elemanları terminal kategorileri olarak adlandırılır. Çünkü bu elemanların her biri giriş tümcesine ait sözcüklerden oluşan birer kümedir.

Bir tümcenin dilbilgisi özellikleri ile tanımlanabilmesi için öbek yapısındaki kelimelerin belirli kurallarla bir araya gelmesi gerekir. Türkçe'nin dil yapısı genel olarak *özne-tümleç-yüklem* üçlemesiyle sıralanmasına rağmen, bunların farklı sıralarda dizilişleri ile de tümce tanımlanabilmektedir. Tümce içerisinde AÖ ve EÖ dışında farklı yapıda sözcükler de bulunabilir [39]. Bunlar kelime ekleri (EKLER), bağlaç ve edat görevini yüklenen diğer bildirimler (Bağlayıcı) olarak sınıflandırılır. Cümle dışı unsurlar (CDU), ünlem grupları, hitaplar, ara sözler ve açıklama cümlelerini içermektedir.

Hemen hemen tüm dillerde olduğu gibi, Türkçe de temel bileşenler olarak özne ve yüklemden oluşur. Diğer dilbilgisi öğeleri içerilmediğinde bile, belirli tümcelerın oluşturulması mümkündür. Tümcelerın bu diziliş Adalı tarafında da ifade edildiği şekilde "Özne  $\lambda$  ...  $\lambda$  Yüklem" eklenmesidir.

ii) Ad Öbeği yapısına ait dilbilgisi kuralları:

$$AÖ \rightarrow AÖ AÖ \mid Ad AÖ \mid Sıfat AÖ$$

$$AÖ \rightarrow Ad EKLER \mid Sıfat EKLER$$

AÖ → Ad | CDU EKLER

AÖ → Bağlayıcı AÖ

şeklinde genelleştirilebilir.

iii) Eylem Öbeği yapısına ait dilbilgisi kuralları:

EÖ → AÖ EÖ

EÖ → Ad EKLER | Eylem EKLER

EÖ → Eylem | BİREylem EKLER

olarak tanımlanır.

Bağlayıcı terminal kategorisine ait küme edat ya da bağlaç içerebilir.

iv) Kelime ekleri birden fazla ek aldığında ekler öbeğine ait kuralları oluşturur ve türetme kuralları

EKLER → Eki EKLER | Eki

şeklinde yazılabilir.

v) Birleşik eylemlere ait dilbilgisi kuralları:

BİREylem → Sıfat EÖ

şeklinde genelleştirilebilir.

vi) Cümle dışı unsurlara ait dil bilgisi kuralları:

CDU → Ünlem AÖ | Sayı

vii) Ad Öbeği içerisinde türetilen sayıların dil bilgisi kuralları:

Sayı → Rakam Sayı | λ

Rakam → {0,1,2,3,4,5,6,7,8,9}

şeklinde rakamların ardışık olarak türetilmesi ile elde edilir.

#### 4.2 Dilbilgisi kurallarından haber başlıklarının çıkarılması

Belirlenen sınırlar çerçevesinde sonlu sayıdaki dilbilgisi kuralından sonsuz sayıda tümcenin sözdizimsel çözümlemesi yapılabilir. Sözdizimsel çözümlemeye örnek vermek üzere TRT Haber sitesindeki belirli haber kategorilerinden başlıklar seçilmiştir. Sözdizimsel analiz gerçekleştirilirken tümcelere ait çözüm ağaçları da oluşturulmuştur. TRT Haber sitesindeki üç farklı başlıktan alınan rastgele seçimlerle “dünya, ekonomi ve spor” kategorilerine ait tümcelerin tanımlanan dilbilgisi kurallarına göre sözdizimsel analizleri aşağıda verilmektedir.

##### 4.2.1 Dünya kategorisi haber başlıkları

Bu kategoride dünya basınından haberler yer almaktadır. Örneğin, “Kiev’de patlama sesleri duyuldu.” haber başlığının sözdizimsel analizine ait

dilbilgisi kuralları aşağıda tanımlanmış ve buradan çözüm ağacı çizilmiştir (Şekil 1).

C → AÖ EÖ

AÖ → Ad AÖ | Ad EKLER

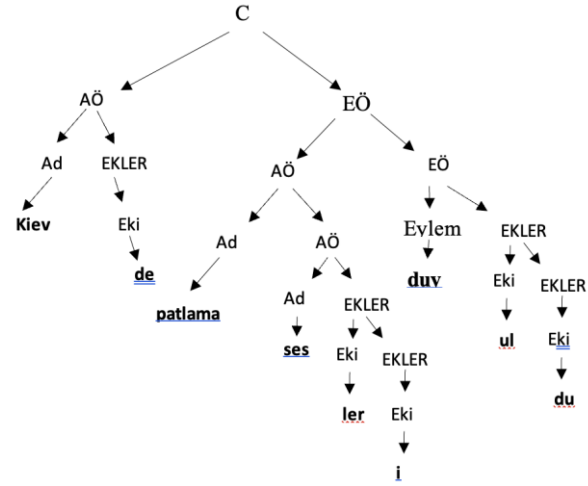
EÖ → Eylem EKLER

EKLER → Eki EKLER | Eki

Ad → {Kiev, patlama, ses}

Eki → {de, ler, i, ul, du}

Eylem → {duy}



Şekil 1. Dünya kategorisine ait bir tümcenin çözümleme ağacı

Haber sitesinin dünya kategorisinden bir başka haber başlığı örneği “Azerbaycan 2023’e coşkuyla girdi.” tümcesidir. Koopman vd. [40] sözcükleri isim, eylem, sıfat, zarf olarak açık sınıf kategorileri ve cümle içerisindeki diğer olası sözcükleri de kapalı sınıf kategorileri şeklinde incelemişler; sayılar da kapalı sınıf kategorilerine dahil edilmiştir. Diğer taraftan Adalı [39] tarafından zaman bildiren öğeler ad öbeği kategorisinde incelenmiştir. Bu nedenle cümle aşağıdaki dilbilgisi kuralları ile türetilmektedir.

Haber başlığının dilbilgisi kurallarından türetilen sözdizimsel analizine ait çözümleme ağacı Şekil 2’ de verilmektedir.

C → AÖ EÖ

AÖ → AÖ AÖ | Ad EKLER | CDU EKLER | Ad

EÖ → AÖ EÖ | Eylem EKLER

EKLER → Eki EKLER | Eki

CDU → Sayı

Sayı → Rakam Sayı | λ

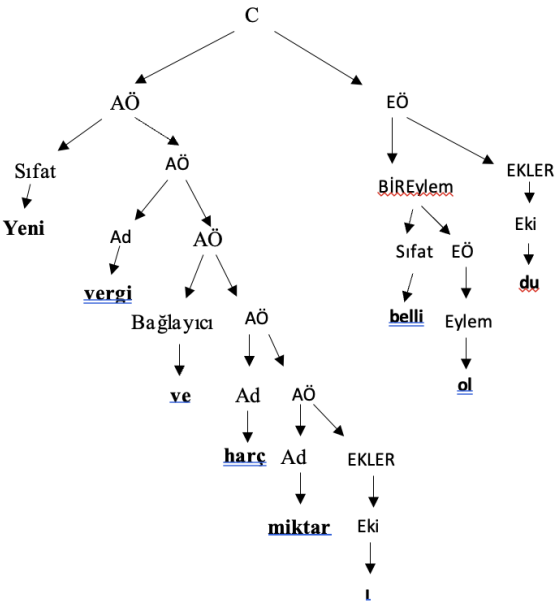




Şekil 4. Ekonomi kategorisine ait bir tümcenin çözümleme ağacı

Ekonomi kategorisinden bir başka haber başlık örneği “Yeni vergi ve harç miktarları belli oldu.” dilbilgisi kurallarından türetilerek sözdizimsel analizi gerçekleştirilmiş ve çözümleme ağacı Şekil 5’te verilmiştir.

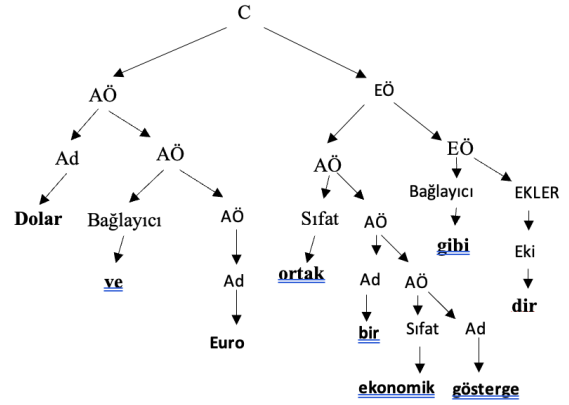
- C → AÖ EÖ  
 AÖ → Sıfat AÖ | Ad AÖ | Bağlayıcı AÖ | Ad EKLER AÖ → Ad  
 EÖ → BİREylem EKLER | Eylem  
 BİREylem → Sıfat EÖ  
 Ad → { vergi, harç, miktar }  
 Eki → { 1, du }  
 Sıfat → { yeni, belli }  
 Bağlayıcılar → { ve }  
 Eylem → { ol }



Şekil 5. Ekonomi kategorisine ait bir diğer tümcenin ağaç görünümü

Ekonomi kategorisinden bir başka haber başlığı örneği “Dolar ve Euro ortak bir ekonomik gösterge gibidir.” ilgili dilbilgisi kurallarından türetilmiş ve sözdizimsel analizi doğrultusunda çözümleme ağacı Şekil 6’daki gibi elde edilmiştir.

- C → AÖ EÖ  
 AÖ → Sıfat AÖ | Ad AÖ | Bağlayıcı AÖ | Ad  
 AÖ → Sıfat Ad  
 EÖ → AÖ EÖ | Bağlayıcı EKLER  
 EKLER → Eki  
 Ad → { Dolar, Euro, gösterge, bir }  
 Eki → { dir }  
 Sıfat → { ortak, ekonomik }  
 Bağlayıcı → { ve, gibi }

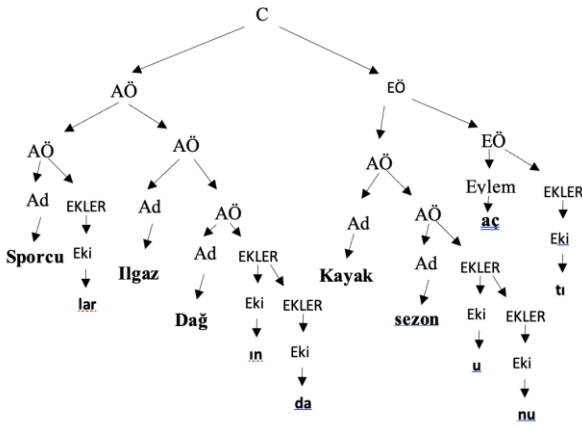


Şekil 6. Ekonomi kategorisine ait bir diğer tümcenin ağaç görünümü

#### 4.2.3 Spor kategorisi haber başlıkları

Bu kategori içerisinde Spora ait haberler yer almaktadır. Örneğin, “Sporcular Ilgaz Dağı’nda kayak sezonunu açtı.” haber başlığının sözdizimsel analizine ait dilbilgisi türetme kuralları oluşturulmuş, çözüm ağacı çizilmiştir (Şekil 7).

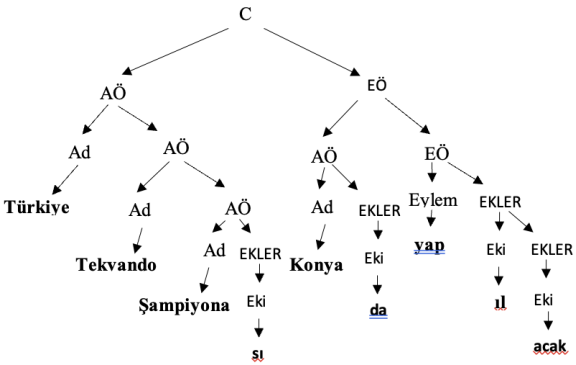
- C → AÖ EÖ  
 AÖ → AÖ AÖ | Ad AÖ | Ad EKLER  
 EÖ → AÖ EÖ | Eylem EKLER  
 EKLER → Eki EKLER | Eki  
 Ad → { sporcu, ilgaz, dağ, kayak, sezon }  
 Eki → { lar, ın, da, u, nu, tı }  
 Eylem → { aç }



Şekil 7. Spor kategorisine ait bir tümcenin çözümleme ağacı

Spor kategorisinden bir başka haber başlık örneği olarak "Türkiye Tekvando Şampiyonası Konya'da yapılacak" seçilmiştir. Bu haber başlığı da sözdizimsel analiz kapsamında incelenip, sözdizimsel karar ağacı Şekil 8'de oluşturulmuştur.

- C → AÖ EÖ  
 AÖ → Ad AÖ | Ad EKLER  
 EÖ → AÖ EÖ | Eylem EKLER  
 EKLER → Eki EKLER | Eki  
 Ad → {Türkiye, Tekvando, Şampiyona, Konya}  
 Eki → { sı, da, ıl, acak }  
 Eylem → { yap }

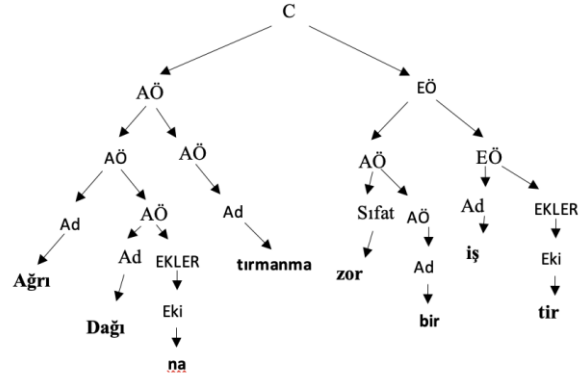


Şekil 8. Spor kategorisine ait bir diğer tümcenin ağaç görünümü

Spor kategorisinden bir başka haber başlık örneği "Ağrı Dağı'na tırmanma zor bir iştir." Bu haber

başlığı da sözdizimsel analiz kapsamında incelenip, sözdizimsel karar ağacı çizilmiştir (Şekil 9).

- C → AÖ EÖ  
 AÖ → AÖ AÖ | Ad AÖ | Ad EKLER | Ad | Sıfat AÖ  
 EÖ → AÖ EÖ | Ad EKLER  
 EKLER → Eki  
 Ad → { Ağrı, Dağı, tırmanma, iş, bir }  
 Eki → { na, da, tir }  
 Sıfat → { zor }



Şekil 9. Spor kategorisine ait bir diğer tümcenin ağaç görünümü

Spor kategorisinden bir başka haber başlık örneği "Filenin Sultanları ilklerin yılı 2023'te üç kupa kazandı, ne büyük bir başarı!" Bu haber başlığı da sözdizimsel analiz kapsamında incelenip, sözdizimsel karar ağacı çizilmiştir (Şekil 10).

- C → AÖ EÖ  
 AÖ → AÖ AÖ | Ad AÖ | Ad EKLER | CDU EKLER  
 AÖ → Sıfat AÖ | Sıfat EKLER | Ad  
 EÖ → AÖ EÖ | EÖ CDU | Eylem EKLER  
 EKLER → Eki EKLER | Eki  
 CDU → Ünlem AÖ | Sayı  
 Sayı → Rakam Sayı | λ  
 Ad → { File, Sultan, yıl, kupa, bir, başarı }  
 Eki → { nin, lar, ı, ler, in, te, dı }  
 Sıfat → { ilk, büyük, üç }  
 Ünlem → { ne }  
 Eylem → { kazan }  
 Rakam → { 0, 2, 3 }



tanımlanması ile tümce çözümlemenin temeline inilmekte ve sözdizimsel analiz kavramı açıklanmaktadır. Türetme kurallarını tanımlayarak haber başlıklarının sözdizimsel çözümlemesine ilişkin oluşturulan bu altyapının benzeri, kod optimizasyonu için de kullanılmaktadır. Böylece gerek kod gerekse sistem hatalarını belirleyen ve düzelten otomatik araçların daha verimli ve hızlı sonuçlar üretmeleri sağlanmıştır.

Sözdizimsel analizin Dil Modellemedeki önemi, tümcenin anlamının doğru tahmin edilmesi olarak özetlenebilir. Tümceler sözdizimsel analiz araçları ile otomatik çözümlendiğinde, uygun anlamsal çıkarımı elde etmek her zaman mümkün olmayabilir. O nedenle incelenen alana özel dilbilgisi kurallarının da eklendiği bir yapı ile çalışmak, tümcenin anlamının daha doğru tahmin edilmesini sağlar. Diğer taraftan biçimbirimsel analiz de anlamın belirlenmesinde önemlidir; bir tümcedeki yüklem çekim ekleri tümcenin anlamını doğrudan değiştirebilir. "Voleybol takımının oyuncuları vize aldılar mı?" ve "Voleybol takımının oyuncuları vize alıyor mu?" tümcelerinde, sadece eylem bildiren sözcüğün çekim ekleri ile tümcelere farklı anlamlar yüklenmektedir.

Sözdizimsel analiz, yapay zekâ geliştirme araştırmalarının temel çalışma öğelerinden biri olmaya her zaman devam edecektir. Zira yapay zekâ teknolojileri geliştikçe sistemlerin yeteneklerini geliştirmek için mevcut yapıdan daha gelişmiş sözdizimsel analiz tekniklerine gereksinim de artacaktır. Bu durumu, problemler karmaşıktıkça daha etkili yazılım geliştirme yöntemlerine olan ihtiyaç ile eşleştirmek mümkündür. Sözdizimsel belirsizliğin daha iyi yönetildiği karmaşık algoritmaların yeni sistemleri daha akıllı yapacağı bir başka gerçektir. Diğer taraftan sözdizimi çalışmalarında farklı bir çözümleme yöntemi olarak, derin öğrenme ve pekiştirmeli öğrenmenin birlikte yürütüldüğü araştırmalar da hızla artmaktadır.

Belirli bir çalışma alanına özel bilgi içeren metinlerin sözdizimsel, semantik, pragmatik yapılarının incelendiği Özellik Mühendisliği ile çözülen problemlerde, geliştirilen sistemin performans ve yetenekleri artmaktadır. Fakat

sözdizimi kullanımı farklı bir bakış açısından değerlendirildiğinde uygulamaya ilişkin etik hususlar ortaya çıkabilmektedir. Metnin sözdizimsel analizi ile kişisel bilginin işlenmesi gizlilik endişelerini arttırmakta ve uygulamalarda sözdiziminin sorumlu olarak ele alınması gündeme gelebilmektedir.

Bir başka değerlendirme ölçütüne göre, gelişmiş sözdizimi modellerinin kullanıcı deneyimini arttırması, uygulamaya olan algıyı da güçlendirmektedir. Örneğin dilin yapısının ve ifadelerin anlamının daha verimli çıkarımı, sanal asistanların daha karmaşık sorgularla daha doğru yanıtlar vermesini sağlamaktadır. Finans sektöründe borsadaki değişimleri daha yüksek doğrulukta tahmin etmede, sözdizimine dayalı duyarlılık analizlerinin etkisi bir başka örnek olarak verilebilir. Sağlık sektöründe, hastalık teşhisinde doktorlara yardımcı olan ve uygun tedavi seçenekleri sunan robotlar sözdizimsel analizleri güçlü yapay zekâ sistemleridir.

Kullanılan uygulamayla uyumlu sözdizimsel analiz tekniğini seçmek önemlidir. Ama sözdizimsel kalıplara fazla güvenmenin sistemde çeşitli sorunlar yaratabileceği unutulmamalıdır. Örneğin, sistemin kısıtlarının dikkate alınmadığı uygulamalarda istenilen performansa ulaşmak zorlaşacaktır. Diğer taraftan performansın sürdürülebilirliği ve değişen gereksinimlere uyum sağlayabilmek için kullanılan modellerde güncellemeler yapmak giderek daha güçlü bir zorunluluk olmaktadır.

### Kaynaklar

- [1] Gece K M. "Türkçe Sözdiziminin Temel Mantığı", Ondokuz Mayıs Üniversitesi Eğitim Fakültesi Dergisi, 8(1)1-97, 1993.
- [2] Chomsky N. "Derivation by phase." In M. Kenstowicz (ed.), Ken Hale: A Life in Language, Cambridge, MA: MIT Press, 1-52, 2001.
- [3] Jager G, Rogers J. "Formal language theory: Refining the Chomsky hierarchy." Philosophical Transactions of The Royal Society B Biological Sciences, 367, 1956-1970, 2012.

- [4] Oflazer K. "Türkçe ve Doğal Dil İşleme." Türkiye Bilişim Vakfı Bilgisayar Bilimleri ve Mühendisliği Dergisi, 5(2), 2016.
- [5] Oflazer K. "Two-level Description of Turkish Morphology." Literary and Linguistic Computing, 9(2), 137-148, 1994.
- [6] Koskenniemi K. "Two-Level Model for Morphological Analysis." Proceedings of the Eighth international joint conference on Artificial intelligence, 2, January, 1983.
- [7] Altan Z. "The Role of Morphological Analysis in Natural Language Processing." Anadolu Üniversitesi Bilim ve Teknoloji Dergisi, 3(1), 56-76, 2002.
- [8] Altan Z. "Dil modellemede Belirsizlik Probleminin Etmenlenmiş Dilbilgisi ile Giderilmesi." In Proceedings of the Turkish Symposium on Artificial Intelligence and Neural Networks, 225-234, 2002-b.
- [9] Dönmez İ, Adalı, E. "Türkçe Hal Ekleri ve Öbekleri Kapsayan Bağlamdan Bağımsız Dil Temsili." Türkiye Bilişim Vakfı Bilgisayar Bilimleri ve Mühendisliği Dergisi, 10(1), 33-40, 2016.
- [10] Bilgin O, Çetinoğlu, Ö, Oflazer, K. "Building a Wordnet for Turkish." Romanian Journal of Information Science and Technology, 7(1-2), 163-172, 2004.
- [11] Fellbaum C. WordNet: An Electronic Lexical Database (Language, Speech, and Communication). Cambridge, MA: MIT Press, 1998.
- [12] Delibaş A. "Doğal Dil İşleme ile Türkçe Yazım Hatalarının Denetlenmesi." Yüksek Lisans Tezi, İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul, 2008.
- [13] Çetinoğlu Ö. "Building a Lexical Functional Grammar for Turkish." Doktora Tezi, Sabancı Üniversitesi, Mühendislik ve Doğa Bilimleri Enstitüsü, İstanbul, 2009.
- [14] Şahin G, Amasyalı M F. "Geniş Metin Koleksiyonlarından İteratif Bilgi Çıkarımı." EMO Bilimsel Dergi, 4(7), 13-20, 2014.
- [15] Agan C, Diri B. "Türkçe Derlemler İçin Söz Dizimsel Görselleştirme ve Sorgulama Aracı." Türkiye Bilişim Vakfı Bilgisayar Bilimleri ve Mühendisliği Dergisi, 9(1), 1-10, 2015.
- [16] Dönmez H G, Becerikli Y. "Türkçe Twitter Verileri Üzerinde Duygu Analizi: LSTM, CNN-LSTM, BERT Algoritmalarının Karşılaştırılması." International Journal of Advanced Natural Sciences and Engineering Researches, 7(6), 132-138, 2023.
- [17] Akın D M, Akın A A. "Türk Dilleri için Açık Kaynaklı Doğal Dil İşleme Kütüphanesi: ZEMBEREK." Elektrik Mühendisliği Dergisi, 431(8), 38-44, 2007.
- [18] Tokcaer S. "Türkçe Metinlerde Duygu Analizi." Journal of Yasar University, 16(63), 1514-1534, 2021.
- [19] İnce E Y. "Web Ortamındaki Yazılı Sınavların Doğal Dil İşleme Yöntemleri ile Değerlendirilmesi." Doktora Tezi, Süleyman Demirel Üniversitesi, Fen Bilimleri Enstitüsü, Isparta, 2016.
- [20] Taşkın S G, Küçüksille E U, Topal K. "Twitter Üzerinde Türkçe Sahte Haber Tespiti." BAUN Fen Bil. Enst. Dergisi, 23(1), 151-172, 2021.
- [21] Gürkan A. "Seçim kampanyalarında yapay zekâ kullanımı seçmene tuzak mı?" [Web sayfası]. Strateji ve Siyaset Araştırmaları Merkezi. Erişim tarihi: 21 Eylül 2023. URL: <https://strasam.org/stratejisiyaset/siyaset-bilimi/secim-kampanyalarinda-yapay-zeka-kullanimi-secmene-tuzak-mi-646>
- [22] Akça T A. "85. Edebi Metinlerde ve Uzmanlık Alan Metinlerinde Makine Çevirisinin Olanakları/Olanaksızlığı: Çevirmenin Değişen Görev Tanımlarına Yeniden Bakmak." Rumeli Dil ve Edebiyat Araştırma Dergisi, (30), 1321-1343, 2022.
- [23] Barut E. "İstatistiksel Makine Çevirisi İle Nöral Makine Çevirisinin Dilbilimsel Parametrelerle Karşılaştırılması: Google Translate." AHBV Akdeniz Havzası ve Afrika Medeniyetleri Dergisi, 4(1), 103-118, 2022.
- [24] Odacıoğlu M C. "Çevrebilimde Yerelleştirme Paradigmasına Doğru." Doktora Tezi, Sakarya Üniversitesi, Sosyal Bilimler Enstitüsü, Sakarya, 2016.
- [25] Yaman İ. "DeepL Translate ve Google Translate Sistemlerinin İngilizce-Türkçe ve Türkçe-İngilizce Çeviri Performanslarının

- Karşılaştırılması." *Söylem Filoloji Dergisi*, 29-41, 2023.
- [26] Wang L, Lyu C, Ji T, Yu Z D. "Document-Level Machine Translation with Large Language Models." *Computer Science*, published: 5 Apr, 2023.
- [27] Çapoğlu N. "Makine Çevirilerinde Sorun Oluşturan Eş Yazımlı Sözcükleri Bulunduran İfadeler için Kural Matrisli Makine Çevirisi (KMMÇ) Modeli." *Bilgisayar Bilimleri ve Mühendisliği Dergisi*, 15 (2), 110, 2022.
- [28] Tantuğ C. "Akraba ve Bitişken Diller Arasında Bilgisayarlı Çeviri için Karma Bir Model." *Doktora Tezi*, İstanbul Teknik Üniversitesi Fen Bilimleri Enstitüsü, İstanbul, 2007.
- [29] Çalış T. "Sözdizimsel Aktarıma Dayalı Makine Çevirisi." *Yüksek Lisans Tezi*, Trakya Üniversitesi Fen Bilimleri Enstitüsü, Edirne, 2017.
- [30] Sadıkov, T. ve Sarıgül, K. "Makine çeviri yöntemleri ve makine çevirisinin bugünkü durumu." *Uluslararası Türkçe Edebiyat Kültür Eğitim Dergisi*, 10(1), 192-205, 2021.
- [31] Microsoft. "Microsoft SwiftKey Keyboard" [Mobil uygulama]. Microsoft Corporation. Erişim tarihi: 27 Ağustos 2023. URL: [https://www.microsoft.com/en-us/swiftkey?activetab=pivot\\_1%3aprimaryr2](https://www.microsoft.com/en-us/swiftkey?activetab=pivot_1%3aprimaryr2)
- [32] Fleksy "Keyboard Technologies" [Mobil uygulama]. Erişim tarihi: 27 Ağustos 2023. URL: <https://www.fleksy.com/>
- [33] Amasyalı M F, Diri B. "Bir Soru Cevaplama Sistemi: BayBilmiş." *Türkiye Bilişim Vakfı Bilgisayar Bilimleri ve Mühendisliği Dergisi*, 1(1), 2006.
- [34] Garanti BBVA "Akıllı Asistan Ugi" [Mobil uygulama]. Garanti BBVA Bankası. Erişim tarihi: 26 Ağustos 2023. URL: <https://www.garantibbva.com.tr/dijital-bankacilik/akilli-asistan-ugi>
- [35] Brill E, Dumais S, Banko M. "An Analysis of the AskMSR Question-Answering System." *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Philadelphia, July, 257-264, 2002.
- [36] Textrazor "Extract Mining from your Text" [Bilgisayar programı]. Erişim tarihi: 26 Ağustos 2023. URL: <https://www.textrazor.com/>
- [37] Summarizebot "Get to know more by reading less!" [Bilgisayar programı]. Erişim tarihi: 26 Ağustos 2023. URL: <https://www.summarizebot.com/>
- [38] Zhang J, Zhao Y, Saleh M, Liu P J. "PEGASUS: Pre-training with Extracted Gap-sentences for Abstractive Summarization." In *Proceedings of the 37th International Conference on Machine Learning, Online Conference*, 13-18 July, 2020.
- [39] Adalı E, Türkçe Doğal Dil İşleme. Akçağ yayınları, Ankara, 2020.
- [40] Koopman H, Sportiche D, Stabler E. "An Introduction to Syntactic Analysis and Theory." Wiley-Blackwell, 2013.