

Çevirmen-Bilgisayar Etkileşiminin Kilit Bileşeni: Doğal Dil İşleme

The Key Element of Translator-Computer Interaction: Natural Language Processing

Araştırma Makalesi / Research Article

Sevda PEKCOŞKUN GÜNER*

* Dr. Öğr. Üyesi, Kırklareli Üniversitesi, Fen-Edebiyat Fakültesi, Mütercim ve Tercümanlık Bölümü, Kırklareli, Türkiye, e-posta: sevda.pekcoskun@gmail.com, ORCID: 0000-0003-2750-3217

ÖZET

Yapay zekânın bir alt alanı olan doğal dil işleme, özellikle son yıllarda gündelik yaşamı etkileyen birçok ürünün ortaya çıkmasını sağlamıştır. İnsan dilini öğrenerek makinenin insanla etkili bir şekilde iletişime geçmesini sağlamak üzere geliştirilen doğal dil işleme teknikleri, yalnızca dilbilimde değil, çeviribilim alanında da paradigma değişikliklerine neden olmuştur. Makine çevirisi sistemleri, otomatik özetleme araçları, terim yönetim sistemleri, metin düzenleme ve düzeltme uygulamaları gibi çeviriye yardımcı araçlar, doğal dil işleme teknikleriyle tasarlanmıştır. Söz konusu araçlar, yapay zekâ alanındaki gelişmelerle paralel olarak sürekli güncellenmekte ve kullanıcı beklentilerine göre iyileştirilmektedir. Doğal dil işlemenin en temel çalışma alanlarından biri makine çevirisidir. Doğal dil işleme tabanlı dil modelleri, çeviri kalitesinin artırılması için sürekli olarak eğitilmekte; bağlama dayalı bilginin sistem işleyişine eklenmesiyle daha başarılı sonuçlar elde edebilmektedir. Bu çalışmanın amacı, doğal dil işleme araçları ve tekniklerinin günümüz çeviri sürecindeki konumunu tartışmaktır. Bu doğrultuda, öncelikle doğal dil işleme kavramından söz edilecek ve doğal dil işleme uygulamalarının çeviri etkinliğini nasıl etkilediği üzerinde durulacaktır. Ardından başta makine çevirisi teknolojileri olmak üzere çevirmenler tarafından kullanılan doğal dil işleme tabanlı uygulamalar, çeviri odaklı bir bakış açısıyla incelenmeye çalışılacaktır.

Anahtar Sözcükler: Yapay Zekâ, Doğal Dil İşleme, Makine Çevirisi, Otomatik Metin Özetleme

ABSTRACT

Natural language processing (NLP), a subfield of artificial intelligence, has led to the development of many products that affect everyday life, especially in recent years. Developed to learn human language, NLP techniques that enable machines to communicate effectively with humans have led to a paradigm shift not only in linguistics, but also in translation studies. Translation aids such as machine translation systems, automatic summarization tools, terminology management systems, and text review and proofreading applications have been designed with NLP techniques. These tools are continuously updated and improved in parallel with user expectations and the developments in the field of artificial intelligence. One of the most fundamental areas of NLP is machine translation. NLP-based language models are continuously trained in order to improve translation quality, and more successful results can be achieved by incorporating contextual information into the system functioning. The aim of this study is to discuss the role of NLP tools and techniques in translating process. To this end, firstly, the concept of NLP will be discussed and how NLP-based applications affect translation activity will be emphasized. Then, NLP-based applications used by translators, especially machine translation technologies, will be analyzed from a translation-oriented perspective.

Keywords: Artificial Intelligence, Natural Language Processing, Machine Translation, Automatic Text Summarization

Giriş

En temel anlamıyla yapay zekâ, insan zekâsı gerektiren öğrenme, anlama, mantıklı düşünme, gerekçelendirme ve sorun çözme gibi bir dizi işlemin makineler tarafından nasıl yerine getirilebileceğini araştıran bir çalışma alanıdır. Günümüzde yapay zekâ çalışmaları önemli bir ivme kazanmış; yapay zekâ ile eğitilen araç ve robotlar gündelik hayatın birer parçası konumuna gelmiştir.

John Rogers Searle (1980), felsefi bir bakış açısıyla yapay zekâyı iki temel kategoride sınıflandırmıştır. Bunlar, zayıf yapay zekâ (İng. *weak AI*) ve güçlü yapay zekâdır (İng. *strong AI*). Zayıf yapay zekâ, insan tarafından gerçekleştirilebilen fiziksel ya da zihinsel işlemlerin yapılmasına ilişkin faydalı ve başarılı uygulamaların ortaya çıkarılabileceğini savunan görüştür. Örneğin, bir montaj hattında çalışan robot kol veya bir resimde yer alan nesnelere ayırt edip betimleyebilen yazılımlar zayıf yapay zekâ yaklaşımı kapsamında değerlendirilebilir. Güçlü yapay zekâ bakış açısına göre ise makineler insanlarla yarışabilecek hatta insanı geçebilecek düzeyde bilişsel yetilere sahip olabilecektir. Yaşadığımız çağda güçlü yapay zekâ yaklaşımıyla tasarlanmış bir ürün ya da yazılımdan söz etmek henüz zor olsa da zayıf yapay zekâ ürünleri gündelik hayatımızda yer almaya; yaşamımızı şekillendirmeye ve değiştirmeye başlamıştır.

Zayıf yapay zekâ yaklaşımıyla ürün tasarımının ardındaki en temel itki, insan zihninin yükünü azaltmaktır. Bazı görevlerini makinelere emanet eden insan, zevk aldığı etkinliklere daha fazla zaman ayırabilir; ayrıca işlerin makinelere devredilmesi zihinsel bir rahatlamayı da beraberinde getirir. İnsan dili ile ilgili görevleri etkin bir şekilde yerine getirmek üzere tasarlanmış doğal dil işleme uygulamaları da zayıf yapay zekâ ürünleri olarak kabul edilmektedir. İnsan dilini anlama ve insanla bilgisayar arasında sağlıklı bir iletişim kurma amacıyla geliştirilen doğal dil işleme, sohbet ve arama motorları gibi gündelik hayatın birer parçası durumuna gelmiş olan birçok yapay zekâ uygulamasının geliştirilmesini sağlar.

Doğal dil işleme yaklaşımları ve ilgili süreçler sonucunda ortaya çıkan ürünler, günümüz çeviri piyasası ve çevirmenlerini de derinden etkilemiştir. Doğal dil işleminin kullanıldığı alanlardan biri makine çevirisidir. Doğal dil işleme teknikleri, kaynak metnin dilbilgisi yapısının ve sözcük dağarcığının belirlenmesi ile analiz edilmesinde yoğun bir şekilde kullanılır. Aynı zamanda erek metin tümcelerinin oluşturulması için üretimsel amaçlı olarak da doğal dil işleme teknikleri uygulanır. Doğal dil işleminin makine çevirisi teknolojileri ile çeviri sürecini doğrudan etkilediği düşünülse de çevirmenin kullandığı birçok aracın tasarımı ve geliştirilmesinde ilgili modellerin payı büyüktür. Masaüstü çeviri yazılımları, otomatik özetleme araçları, terminoloji yönetim sistemleri, çeviri bellekleri ve yazım denetleyiciler artalanlarında doğal dil işleme yöntemlerini kullanan ve çevirmenin iş istasyonunda yerini almış olan yazılımlardır. Çevirmenler, doğal dil işleme yazılımlarını daha kısa sürede çeviri yapmak, daha az stresli bir çeviri süreci geçirmek ve daha kaliteli bir erek metin üretmek amacıyla kullanırlar. Doğal dil işleme modelleri kullanılarak tasarlanmış çeviriye yardımcı araçlar, çevirmenin daha hızlı çeviri yapmasını sağlayabildiği gibi; özellikle teknik alanda farklı çeviri çözümleri de sunabilir. Örneğin, terim yönetim araçları çevirmenin tüm çeviri süreci boyunca terim tutarlılığını sağlamasına yardımcı olur. Metin kalitesine yönelik değerlendirme araçları ise erek metnin okunabilirliği ile ilgili çözümler sunar. Dolayısıyla, doğal dil işleme

teknikleriyle geliştirilmiş ürünler çeviri sürecine doğrudan ya da dolaylı olarak katkı sunabilmektedir.

Hem genel kullanıcı hem de profesyonel kullanıcılara hitap eden makine çevirisi sistemleri, doğal dil işleme tabanlı modellerin en popüleridir. Bu çalışmada, özellikle çevirmenler tarafından kullanılacak doğal dil işleme uygulamaları örneklerle betimlenmeye çalışılmıştır. Çalışmanın temel amaçlarından biri, günümüzde yaygın şekilde kullanılan doğal dil işleme araçları ve yöntemlerinin çeviri odaklı bir bakış açısıyla incelenmesidir. Çalışmanın diğer bir amacı ise, söz konusu doğal dil işleme araçlarının çeviri sürecindeki konumu ve işlevselliğinin sorgulanmasıdır.

Çeviribilim alanyazınında görece yeni bir kavram olan doğal dil işleme konusu çalışmanın ilk bölümünde detaylı şekilde ele alınacaktır. İkinci bölümde makine çevirisi kavramı ve makine çevirisi türleri yazılı çeviri uygulamaları bağlamında incelenecektir. Üçüncü bölümde, makine destekli sözlü çeviri kavramı ve otomatik konuşma tanıma yaklaşımları irdelenecektir. Çalışmanın dördüncü bölümünde ise doğal dil işleme tabanlı dijital araçlar kullanılarak gerçekleştirilebilen otomatik metin özetleme ve özetleyerek çeviri işlemleri ele alınacaktır. Ayrıca çalışmada, doğal dil işleme ve çeviribilim arasındaki ilişki vurgulanarak, ilgili araç ve yöntemlerinin çeviri sürecine nasıl entegre edilebileceği betimlenmeye çalışılacaktır.

1. Doğal Dil İşleme Kavramına Genel Bir Bakış

Doğal dil işleme (İng. *Natural Language Processing, NLP*) konusuna değinmeden önce, “doğal dil” kavramını tanımlamamız faydalı olacaktır. Doğal dil, bilgisayarlar tarafından kullanılan yapay diller veya programlama dillerinin aksine, insanların birbirleriyle iletişim kurmak için kullandıkları dil anlamına gelir. Doğal diller, gündelik konuşmalarda kullanılan sözcükler ve söz öbeklerini, dilbilgisi yapılarını ve dilsel örüntüleri içerir. Sören Stenlund (1990, ss. 5-6), doğal dil kavramının, “formel dil”¹ ve “yapay dil” ifadelerine karşıt olarak kullanıldığını ifade eder. Doğal dilin oluşum süreci, yapay ya da yapılandırılmış bir dilin oluşturulma adımlarından oldukça farklıdır ve doğal diller “formel diller” kategorisine girmez. Ancak, uygulamalı alanda doğal bir dil, sanki formel bir dilmiş gibi ele alınır ve inceleme malzemesi olarak kullanılır.

Doğal diller ve formel diller arasındaki farkları özetlememiz gerekirse, doğal bir dilin dizgesel olduğunu; biçim bilgisi, sözdizim ve anlam bilgisi gibi katmanların birbiriyle etkileşimi sonucu ortaya çıktığını söyleyebiliriz. Yapay ya da formel diller ise, sözdizimi ve dilbilgisi açısından sınırları iyi belirlenmiş kurallar yardımıyla “kesin ve net” bileşenler içerecek şekilde tasarlanır (Scott, 2015). Doğal diller, herhangi bir planlama olmadan,

¹ Burada “formel dil” resmi dil ya da alana özgü dil anlamında kullanılmamaktadır. Bilgisayar bilimlerinde “formel dil”, bilgisayarlar, makine öğrenimi algoritmaları ve ilgili yapay zekâ sistemleri tarafından kullanılmak üzere tasarlanmış bir dili ifade eder. Böylece makineler, bir dizi katı kural yardımıyla, doğal dili (insan dili) işleyebilir, anlayabilir ve doğal dille etkileşim içine girebilir. Formel diller, programlama dillerini, biçimlendirme dillerini ve ontolojileri içerebilir. Formel dillerin kullanımı, insanlar ve makineler arasında tutarlı bir iletişim kurulmasını sağlayarak, karmaşık sistemlerin ve uygulamaların tasarlanabilmesini mümkün kılar. Formel diller, matematiksel kesinliğe dayalı ve içerisinde doğal dillerdeki gibi anlam bulanıklığı barındırmayan dillerdir.

rastlantısal olarak oluşur. Ancak, bu dillerin kendine özgü kuralları da vardır, örneğin belli bir sözcüğün bir kavramı anlamsal olarak karşılması için doğal diller belirli ilkelere uyarlar. Başlangıçta söz konusu sözcüğün, belirtilen kavrama atanmış olmasının hiçbir nedeni yoktur. Doğal dillerin aksine, formel diller, belirli kavramları ifade etmek için kullanılan ve özel olarak geliştirilmiş terimlerden oluşan sözcük dağarcığına sahiptir. Bu nedenle, formel dillerin ifade gücü sınırlıdır ve yalnızca belirli kavramlarla görevleri ifade etmek için kullanılırlar. Doğal dillerde, aynı bilgi farklı şekillerde ifade edilebilir; bu açıdan doğal diller konuşucularına çok sayıda yapı ve sözcük sunar. Tüm doğal diller zaman içinde gelişir ve değişir, söz konusu değişimlerin de farklı nedenleri ve yöntemleri vardır (Millward ve Hayes, 2012). Yapay diller ise değişmez yapıda olup, bilgisayarlar tarafından kolayca işlenebilecek şekilde tasarlandığı için, insan ve makine arasında daha verimli iletişim kurulmasına olanak tanır.

Doğal dil işleme, bilgisayar bilimleri, dilbilim ve yapay zekânın kesişim kümesini oluşturan, disiplinler arası yapıya sahip bir çalışma alanıdır. Doğal dil verisinin işlenmesi ve analiz edilmesi amacıyla bilgisayarın nasıl programlanabileceğini inceleyen bir disiplin olan doğal dil işlemenin nihai hedefi, doğal dilin tüm bileşenleriyle bilgisayar tarafından anlaşılabilmesini sağlamaktır. Doğal dil işlemenin kullanıcının işini kolaylaştırmak ve bilgisayarla doğal dilde iletişim kurma isteğine yanıt vermek için ortaya çıktığı da söylenebilir.

Bir dil, bir dizi kural veya sembol kümesi olarak tanımlanabilir. Semboller bir araya getirilir ve bilgiyi aktarmak için kullanılır. Bu semboller, kurallar tarafından yönetilir. Doğal dil işleme süreçleri temel olarak “doğal dilin anlaşılması” ve “doğal dilin yeniden üretilmesi” olmak üzere iki kısma ayrılmaktadır (Khurana, vd., 2022). “Anlama” aşaması dilbilimsel bileşenlerden oluşmakta olup, sesbilgisi, biçimbilgisi, sözdizim, anlambilgisi ve edimbilgisi konularını kapsar. “Yeniden üretme” adımı ise doğal dilde bir metnin oluşturulmasını içerir. Bu bağlamda, doğal dil işleme alanındaki araştırmalar toplamda beş kategoride sınıflandırılabilir: (1) Doğal dilin anlaşılması, (2) Doğal dilin yeniden üretilmesi, (3) Konuşma ya da ses tanıma, (4) Makine çevirisi, (5) Yazım ve dilbilgisi denetimi (Joseph, vd., 2016, s. 208).

Makine öğrenmesi ve derin öğrenme alanlarındaki gelişmeler, doğal dil işlemeyi olumlu yönde etkilemiş; genel dil modelleri oluşturmak için büyük miktarlarda dilsel veriden yararlanılmaya başlanmıştır. Bu tür modeller, çeşitli görevleri yerine getirmek için kullanılabilen, otomatik konuşma tanıma ve makine çevirisi gibi daha üst düzey işlemlere entegre edilebilen, bir dilin matematiksel temsilleridir (Fantinuoli, 2023). Dil modellerinin kullanımına yönelik olarak verilebilecek en başarılı örneklerden biri ise son dönemde tüm dünyada popüler olan *ChatGPT*²’dir. *ChatGPT*, artalanında *GPT-3* ve *GPT-4* dil modellerini kullanan, çevrimiçi erişilebilen ve sohbet için özelleştirilmiş olan bir yazılımdır. Bu yazılım, doğal dilde verilen kullanıcı girdilerine, yine doğal dilde yanıtlar üretebilmekte ve böylece bir diyalog ortamının simüle edilmesini sağlayabilmektedir.

Doğal dil işleme, günümüzde milyarlarca dolarlık bir endüstri haline gelmiştir. Doğal dil işleme çalışmaları, sağlıktan satış ve pazarlamaya kadar birçok sektörde devrim yaratma

² <https://chat.openai.com> (Erişim tarihi: 20.05.2023)

potansiyeline sahiptir. Doğal dil işleme, aynı zamanda sıradan bir insanın her gün kullandığı birçok dijital uygulamanın da bel kemiğini oluşturmaktadır. *Microsoft Word* ve *Libre Office* gibi sözcük işlemcilerden *Google* ve *Yandex* gibi arama motorlarına kadar birçok aracın artalanında doğal dil işleme mekanizmaları bulunmaktadır.

Doğal dil işleme yöntemleri kullanılarak oluşturulmuş ve yaygın kullanım bulmuş olan bazı uygulamalar; dillerarası çeviri sistemleri, sanal asistanlar³, arama motorları, yapay zekâ tabanlı yazma asistanları ile yazım denetim araçları, otomatik tamamlama ve düzeltme araçları, sohbet robotları, otomatik içerik düzenleme sistemleri ve konuşmanın metne dönüştürülmesini sağlayan araçlardır (Madill, 2022). Söz konusu uygulamalar, günümüzde oldukça yaygın kullanım bulmuş; dijital çağın gereklerine ve kullanıcıların beklentilerine göre yeniden şekillenmiş ve güncellenmiştir.

Doğal dil işleme kullanılarak geliştirilen araçlar bir yandan bireylerin gündelik yaşamlarının bileşenleri haline gelmiş; diğer yandan özellikle dilbilim ve çeviribilim gibi dille ilgili uzmanlık alanlarında çalışan araştırmacılar için hem birer inceleme nesnesi hem de farklı metin türlerinde makine çevirisinin başarımının ölçülmesi gibi belirli konularda gerçekleştirilen araştırmalarda kullanılan yardımcı araçlar olarak konumlanmıştır.

Kısaca ifade etmek gerekirse, insan dilinin bilgisayarlar tarafından işlenmesini içeren tüm süreçler doğal dil işleme kapsamına girmektedir. Bu nedenle çeviri ile ilgili tam ya da yarı otomatik sistemlerin de doğal dil işleme çalışmaları ile doğrudan ilişkisi vardır. Doğal dil işleme ve çeviribilim, insan dilini; başka bir deyişle doğal dili kavrama ve anlamlandırma konusunda aynı amacı paylaşan birbiriyle ilişkili iki alandır. Doğal dil işleme, dili inceleme ve yorumlama amaçlı algoritmalar geliştirme gibi dilin hesaplamalı yönlerine odaklanırken, çeviri etkinliği kaynak dilde oluşturulmuş içeriğin bilgi kaybı olmaksızın, erek dile aktarımına yoğunlaşır. Doğal dil işleme, otomatik çeviri sistemleri geliştirmek ve mevcut çeviri sistemlerinin doğruluğunu ve verimliliğini artırmak için kullanıldığından, çeviribilim açısından oldukça önemli bir yere sahiptir. Sonraki bölümlerde doğal dil işleme yöntemleri kullanılarak oluşturulmuş ve çevirmenin iş istasyonunda yerini almış olan bileşenlerden söz edilecektir.

2. Çeviri Süreçlerinde Kullanılan Doğal Dil İşleme Tabanlı Sistemler

2.1. Makine Çevirisi Sistemleri

Bilindiği gibi makine çevirisi, algoritmalar kullanarak metni bir dilden diğerine otomatik şekilde çevirme işlemi anlamına gelmektedir. Doğal dil işleme ise bilgisayarların insan dilini anlamasını, analiz etmesini ve üretmesini sağlar; bu nedenle makine çevirisi sistemlerinin geliştirilmesinde oldukça önemli bir rol oynar. Makine çevirisinin, doğal dil işlemenin en zorlu, en temel ve en gerekli görevi olduğu birçok araştırmacı tarafından özellikle belirtilmiştir (Manning ve Schütze, 1999, s. 463; Ma ve Tang, 2021, s. 211; Rothman, 2021, s. 343). Doğal

³ Doğal dille verilen komutlara uygun işlevleri yerine getiren yazılım araçlarıdır. Örneğin, *Google Asistan*'a sorular sorulabilir ve ondan belli bir görevi yerine getirmesi beklenebilir. *Google Asistan* şu anda Türkçenin de içinde bulunduğu 30 dilde sesli komutları algılayabilmektedir.

dil işleme, makine çevirisi sürecinde sisteme yüklenen kaynak metni; sözcükler, söz öbekleri ve tümceler gibi daha küçük bileşenlere ayırmak ve ardından metnin anlamını belirleyebilmek amacıyla bu bileşenleri analiz etmek için kullanılır. Bu süreç, konuşma parçası etiketleme, ayırıştırma, anlamsal analiz ve duygu analizi gibi çeşitli alt görevleri içerir.

Makine çevirisi sistemlerinde bir kaynak metni, belirlenen erek dile aktaran modeller oluşturmak için de doğal dil işleme teknikleri kullanılır. Bu modeller, dil örüntülerini ve olası çeviri birimlerini belirlemek için kural tabanlı yaklaşımları, istatistiksel yöntemleri ve sınır ağlarını kullanabilir. Paralel metinlerden oluşan büyük derlemelerle eğitilen modeller, farklı dil çiftlerinde birbirleriyle hizalanmış kaynak ve erek metin birimlerini içerir.

Doğal dil işleme aynı zamanda makine çevirisi çıktısı üzerinde düzeltelerin yapılmasına ilişkin aşamayı oluşturan art düzenleme (İng. *post-editing*) sürecinde de görev alabilir. Art düzenleme, erek metindeki hataların belirlenmesini ve düzeltilmesini kapsar; dolayısıyla, insan çevirmenlere yardımcı olmak amacıyla, art düzenleme sürecinde doğal dil işleme teknikleri kullanılabilir. Örneğin, bir makine çevirisi sistemi dilbilgisi açısından doğru; ancak anlamsal açıdan sorunlu olan bir tümce üretebilir. Bu tür durumlarda doğal dil işleme, hataların belirlenmesine ve düzeltilmesine yardımcı olur. Kısaca ifade etmek gerekirse, bilgisayarların insan dilini anlamasını ve üretmesini sağlayan doğal dil işleme kapsamında, makine tarafından üretilen erek metinlerin kalitesini ve doğruluğunu artırmaya yarayan yöntemler de geliştirilmektedir.

Makine çevirisi sistemlerinin kullanıcıları temelde iki kategoriye ayrılabilir:

- **Genel Kullanıcılar:** Belirli bir konuda mesleki ve uzmanlık bilgisi olmayan, toplumun farklı ekonomik ve sosyal kesimlerinden gelen, eğitim düzeyleri birbirlerinden farklı olan geniş bir kitleyi temsil eden kullanıcılardır. Bu kullanıcılar aynı zamanda “ağ toplumu”nun birer ferdidir. Manuel Castells (2004, s. 3) ağ toplumunu, “mikro-elektronik tabanlı bilgi ve iletişim teknolojilerini temel alan ağlardan oluşan toplum” olarak tanımlamıştır. Bir başka deyişle, “ağ toplumu”, iş, iletişim ve yönetim amaçları doğrultusunda küresel ağları kullanan bir toplumu ifade eder (Anderson, 2014, s. 156). Çoğunlukla ücretsiz olarak kullanılabilen çevrimiçi makine çevirisi sistemleri, ağ toplumu tarafından kullanılmakta, bireyler bilmedikleri bir dilde yazılmış metinlerin ne demek istediğini böylece anlayabilmektedir. İnternet erişimi olan tablet, akıllı telefon, bilgisayar gibi teknolojik araçlara sahip kullanıcılar makine çevirisi sistemlerinden rahatlıkla yararlanabilmektedir. Son dönemde geliştirilen mobil uygulamalarla, kaynak metni ilgili platforma yazmak ya da kopyalamak gerekmeden, akıllı telefonun doğrudan kaynak metin üzerine tutulmasıyla da çeviri gerçekleştirilebilmektedir. Ayrıca, makine çevirisi sistemleri içerdikleri ses tanıma özelliği ile kaynak dilde seslendirilen tümceleri yazılı ve sözlü olarak erek dile aktarabilmektedir. Sonuç olarak, dijital okuryazar olan ve gerekli teknik ekipmana sahip tüm kullanıcılar makine çevirisi sistemlerini amaçları doğrultusunda verimli bir şekilde kullanabilir. Bu sistemler, genel kullanıcılar tarafından çoğunlukla yabancı dildeki metnin “ne demek istediğini” anlamak için kullanılır. Bu nedenle, genel kullanıcının makine çevirisi sistemlerinden beklentisi

kaliteli bir çeviriden ziyade, metnin içeriği hakkında bilgi sahibi olmalarını sağlayabilecek yeterli bir çeviridir.

- **Uzman Kullanıcılar:** Bu kategorideki kullanıcılar, mesleki ve teknik yeterliğe sahip, belirli uzmanlıkları bulunan ve özel alan bilgisi olan kişilerdir. Makine çevirisi sistemlerini kullanım amaçlarına göre, uzman kullanıcıları gruplara ayırmak mümkündür. Bu sistemleri yaygın olarak kullanan uzman kullanıcıların başında profesyonel çevirmenler ve çeviribilim öğrencileri gelmektedir. Günümüzde çevrimiçi makine çevirisi sistemleri, masaüstü ve bulut tabanlı yardımcı araçlara kıyasla daha fazla kullanılmaktadır. Bu durumun temel nedenleri ise, makine çevirisi sistemlerinde bir çeviri belleğine gereksinim duyulmaması, sistemlerin otomatik çeviri yapması, sistemden oldukça hızlı sonuçlar alınabilmesi ve çoğu makine çevirisi yazılımının ücretsiz olmasıdır. Daha fazla miktarda dilsel materyali daha kısa sürede çevirme olanağı sunan sistemler, kullanıcı dostu arayüzleriyle çevirmenin iş istasyonunda yerini almıştır. Çevirmenler yanında, makine çevirisi sistemlerinin geliştirilmesinde görev alan hesaplamalı dilbilim uzmanları, bilgisayar ve yazılım mühendisleri ile bilişim alanında çalışmalar yapan uzmanlar bu sistemleri etkin şekilde kullanmaktadır. Aynı zamanda, tıp, hukuk, mühendislik gibi özel alanlarının mensubu olan profesyoneller de makine çevirisi sistemlerinden yararlanmaktadır. Örneğin, muayene esnasında bir doktor, dil engeli ile karşı karşıya kalabilir ve ortamda profesyonel bir tercüman da yoksa, anamnez almak ve hasta öyküsünü netleştirmek, klinik bir tanıyı gözden geçirmek veya önerilen tedavi planını yeniden ifade ederek, hastanın süreci anlamasını kolaylaştırmak için makine çevirisi sistemleri kullanılabilir (Randhawa, vd., 2013, ss. 382-383). Doktorlar, ayrıca hastaları makine çevirisi sistemine yönlendirerek, sistem aracılığıyla soru sormaya veya sorulara yanıt vermeye teşvik edebilir. Uzman kullanıcılar, genel kullanıcılardan farklı olarak, makine çevirisi sisteminden kaliteli, işlevsel ve kabul edilebilir bir çeviri beklerler. Bu nedenle, uzman kullanıcıların yer aldığı makine çevirisi süreçlerinde ön düzenleme ve art düzenleme gereksinimi daha fazladır.

Makine çevirisi sistemleri sürekli olarak güncellenmekte ve geliştirilmektedir. Bu yazılımların pazar payı 2020 yılında 650 milyon doları bulmuştur. 2027 yılına gelindiğinde ise, makine çevirisi pazarının 3 milyar doları aşacağı öngörülmektedir⁴. Tüm yatırımların nihai amacı, makine çevirisi sistemlerinin daha kaliteli erek metinler üretebilmesidir. Sistemler, yeni ve daha kapsamlı özelliklerle donatıldıkça, insan eliyle gerçekleştirilen düzeltilere gereksinim de azalmaktadır. Makine çevirisinin tarihsel gelişimine bakıldığında, başlangıç noktası (1950ler) ile günümüz arasında çıktının başarımı açısından oldukça büyük bir fark bulunduğu açıktır. Doğal dil işleme yöntem ve modellerinin gelişimine paralel olarak makine çevirisi sistemleri evrilmiş ve gelişmiş, çıktılar insan tarafından üretilen çevirilerin başarımına yaklaşmıştır. Günümüzde doğal dil işlemenin bir alt alanı olarak kabul edilen makine

⁴ <https://www.transperfect.com/blog/evolving-machine-translation-and-translation-memory-landscape-life-sciences> (Erişim tarihi: 20.05.2023)

çevirisinin tüm türlerinde farklı teknikler kullanılmaktadır. Makine çevirisi türlerini aşağıdaki şekilde açıklayabiliriz:

2. 1. 1. Kural Tabanlı Makine Çevirisi Sistemleri

Kural tabanlı makine çevirisi (İng. *Rule-Based MT*), en ilkel mod olarak kabul edilir. 1950li yıllardan 1980lere kadar kullanılan makine çevirisi sistemlerinin oluşturulmasında kullanılan yöntemler kural tabanlıdır. Kural tabanlı makine çevirisinde, kaynak ve erek dilin dilbilgisi yapıları dikkate alınarak çeviri yapılır. Bu sistemlere kural tabanlı denilmesinin nedeni ise, “kaynak dilden erek dile çeviri yapılabilmesini mümkün kılmak için manuel olarak geliştirilen bir dizi kural aracılığıyla yeniden yapılandırma çerçevesinin oluşturulmasıdır” (All Languages Ltd., 2016, s. 5). Sistemlerin arka planında yazım denetleyiciler, tek dilli, çift dilli sözlükler ve sözlükçeler bulunmaktadır. Kural tabanlı makine çevirisi, büyük oranda ilgili dil çiftlerine özgü sözlüklere bağlı olarak çalıştığından, çıktıda sözdizimi açısından ve anlamsal düzeyde hatalar ortaya çıkabilir. Bu nedenle, kural tabanlı makine çevirisi sistemlerinden kaliteli bir çeviri elde edilmesi yoğun bir düzeltmenlik çabasına bağlıdır. Belli dil çiftleri ve özel alanlara yönelik olarak tasarlanmış farklı araçlar da sistemlere entegre edilebilir. Ancak, bir dil çifti ya da alan için tasarlanmış bir araç başka dil çiftleri ve özel alanlarda işe yaramayacaktır.

Kural tabanlı sistemlerde, temelde altı bileşen bulunur. Kaynak dil tabanlı morfolojik çözümleyici; kaynak dildeki sözcükleri çözümler ve bu sözcüklere ilişkin morfolojik veriler sunar. Kaynak dil tabanlı ayrıştırıcı ise, kaynak dil tümcelerini analiz eden sözdizimi çözümleyicisidir. Çevirici, kaynak dildeki sözcüğü erek dile aktarır. Erek dil tabanlı morfolojik üretici, kendisine sunulan dilbilgisi veri tabanını kullanarak uygun erek dil bileşenlerini üretir. Erek dil tabanlı ayrıştırıcı, uygun bulunan erek dil tümcelerinin derlenmesinden sorumludur. Son olarak, sözlükler kullanılır. Kural tabanlı sistemlerin artalanında kullanılan sözlükleri üçe ayırmak mümkündür. Morfolojik analizin başarılı bir şekilde gerçekleştirilmesi için kaynak dile özgü tek dilli bir sözlük, kaynak dil tabanlı morfolojik çözümleyici tarafından kullanılır. Çift dilli bir sözlük, kaynak dildeki sözcükleri erek dile aktaran çevirici tarafından kullanılır. Erek dile özgü tek dilli sözlük ise, erek dilde doğru sözcüklerin üretilebilmesi için erek dil tabanlı morfolojik üretici tarafından kullanılır (Hettige ve Karunananda, 2011).

Dilbilimsel kesinlikte tasarlanmış olmalarına rağmen, kural tabanlı makine çevirisi sistemlerinin bazı kısıtları mevcuttur. Öncelikle, çevrilmesi gereken metin miktarı arttıkça ve farklı özel alanlarda metinlerin çevrilmesi gerektiğinde, artalanında çalışan sözlükler yetersiz kalabilmektedir; yeni ve kapsamlı sözlüklerin tasarlanması ise hem zahmetli hem de masraflı bir iştir. Dilbilgisi verilerinin sistemlere çoğunlukla manuel olarak girilmesi gerekmektedir, bu durum da dilbilimciler üzerinde büyük bir iş yükü oluşturmaktadır. Aynı zamanda, belli metin türlerine yönelik dilbilgisi ve sözcük tabanlı kaynakların oluşturulması uzun bir araştırma sürecini gerektirmektedir. Kural tabanlı bir sistemde kültüre özgü unsurların, deyimlerin ya da metaforik ifadelerin çevrilmesi zordur. Bağlamdan ziyade, tekil olarak sözcüklerin çevirisi ön plandadır. Bu nedenle, anlamsal düzlemde kaliteli makine çevirisi çıktıları kural tabanlı sistemlerle üretilememektedir.

2. 1. 2. Örnek Tabanlı Makine Çevirisi Sistemleri

Örnek tabanlı makine çevirisi (İng. *Example-Based MT*) sistemleri genellikle 1980li yıllardan 1990lara kadar kullanılmıştır. Bu yöntemde sistem, karşılaştığı yeni tümceleri daha önce çevrilmiş örneklere benzeterek çevirmeyi öğrenir. Örnek tabanlı yaklaşım, kaynak dildeki her bir tümcenin erek dildeki çevirisiyle eşleştirildiği iki dilli tümce çiftlerinden oluşan bir veri tabanına dayanır. Sistem, yeni tümcelerin nasıl çevrileceğini öğrenmek için veri tabanındaki tümce çiftlerini bir başvuru derlemi olarak kullanır. Bu yönüyle örnek tabanlı sistemler, çeviri belleklerini andırmaktadır.

Kaynak dilde çevrilmesi gereken yeni bir tümceyle karşılaştığında sistem, veri tabanında yer alan benzer tümceleri arar ve bunları uygun bir çeviri üretmek için kullanır. Başka bir deyişle, kaynak tümceyi veri tabanındaki her bir tümce ile benzerlik puanlarına göre karşılaştırarak, en benzer tümceleri seçer. Sistem sonraki adımda ise, veri tabanında yer alan en benzer kaynak tümcenin çevirisini seçer ve bu çeviriyi çevrilmesi istenen tümceyle uyumlu hale getirir.

Kural tabanlı sistemlerle karşılaştırıldığında daha başarılı sonuçlar vermesine rağmen, örnek tabanlı makine çevirisinin bazı kısıtları bulunmaktadır. Öncelikle, sistemler arka plandaki veri tabanına tam anlamıyla bağımlıdır ve veri tabanının kalitesi ile boyutu sistemin çeviri kalitesini önemli ölçüde etkiler. Veri tabanı küçükse veya çevrilecek metin türünden az sayıda örnek içeriyorsa, sistemin başarımı düşebilir. Ayrıca çalışma sistemi, benzetim (analoji) yöntemine dayalı olduğundan, yeni ifadeler ya da karmaşık tümce yapıları ile karşılaştığında anlamsal açıdan sorunlu makine çevirisi çıktıkları oluşturabilir.

2. 1. 3. İstatistiksel / Derlem Tabanlı Makine Çevirisi Sistemleri

1990lı yıllardan itibaren istatistiksel / derlem tabanlı makine çevirisi (İng. *Statistical / Corpus-Based MT*) yöntemleri denenmeye başlanmıştır. Bu yöntemler, doğal dilin çevirisi sürecine “makine öğrenmesi sorunu” olarak yaklaşır. Güçlü bir bilgisayar, istatistiksel olarak oluşturulmuş mantığı nasıl uygulayacağını öğrenebilir, böylece ilgili dildeki bir dizi kaynak metin, erek dilin yapısal ve anlamsal özelliklerine göre tasarlanmış bir dizi sözcüğe dönüştürülebilir (All Languages Ltd, 2016). İstatistikler, maksimum olasılık modellerine dayanır, istatistiksel modeller ise tek ve çift dilli metin derlemlerinden elde edilen verilerle oluşturulur. Kural tabanlı sistemlerde manuel olarak girilmesi gereken birçok dilsel veri varken; istatistik tabanlı sistemler, öğrenmelerini derlemler üzerinden gerçekleştirir. Makine tarafından okunabilen ve işaretlenebilen birçok paralel derlem bulunmaktadır. Paralel metin yığınları kullanılarak dilsel örüntüler belirlenmeye çalışılır, çeviri çözümlerine de istatistiksel olarak ulaşılır. İstatistiksel makine çevirisi modellerinin eğitilmesinde gerçek metinlerden oluşan paralel derlemler kullanılır.

Avrupa Parlamentosu’na ait belgelerden oluşan *Europarl* Derlemi paralel derlemlere örnek olarak verilebilir. Bu derlem, 1996 yılından beri hazırlanan Avrupa Parlamentosu tutanaklarını içerir. Derlem, Avrupa Birliği’nin 21 resmi dilinde (Fransızca, İtalyanca, İspanyolca, Portekizce, Rumence, İngilizce, Felemenkçe, Almanca, Danca, İsveççe, Bulgarca, Çekçe, Lehçe, Slovakça, Slovence, Fince, Macarca, Estonca, Letonca, Litvanca ve Yunanca),

her bir dil için 30 milyon sözcüğü aşan dilsel veri sunar. *Europarl*, istatistiksel makine çevirisi sistemlerinin geliştirilmesi için yoğun bir şekilde kullanılmıştır. Philipp Koehn, istatistiksel makine çevirisi sistemlerinde kullanılacak bir paralel derlemin tasarım sürecini beş adımda açıklar: (1) ham verilerin elde edilmesi (örneğin genel ağ taramaları), (2) paralel metin bölümlerinin ayıklanması ve eşleştirilmesi (belge hizalama aşaması), (3) metnin tümcelere ayrılması (tümce bölümlenme), (4) istatistik tabanlı makine çevirisi sistemi için derlem hazırlanması (normalleştirme, bölütlenme), (5) bir dildeki tümcelerin, erek dil karşılıklarıyla hizalanması (2005, ss. 79-86).

İstatistiksel makine çevirisi sistemleri, tek dilli ve çift dilli derlemlerin analizini temel alarak çalışır. Analizler sonucunda, çevirinin gerçekleştirilmesi için istatistiksel modeller oluşturularak, en uygun kaynak dil karşılığının sistem tarafından kullanıcıya sunulması amacıyla istatistiksel ağırlıklar hesaplanır. Gerçek paralel metinlerle çalışıldığından, makine bağlam bilgisini de belli ölçüde öğrenir. Bu nedenle, bu sistemler, kural tabanlı sistemlere göre daha akıcı ve anlama dayalı makine çevirisi çıktıları üretebilir. Bu da düzeltmenlik çabasını azaltmaktadır. 2.1.2’de değinilen örnek tabanlı makine çevirisi sistemleri ile istatistiksel makine çevirisi sistemleri arasındaki temel fark ise, örnek tabanlı sistemlerin çeviri örneklerinden oluşan bir veri tabanına dayanması, istatistiksel makine çevirisinin ise çevirileri oluşturmak için mevcut çeviri örneklerinden çıkardığı istatistiksel modelleri kullanmasıdır. Olasılık tabanlı çalıştığından, istatistiksel makine çevirisi hatalı da olsa mutlaka bir çeviri çözümü sunabilirken; örnek tabanlı sistemler art alanlarında yer alan derlemlerde kaynak metne benzer yapılar mevcut değilse, herhangi bir çeviri çözümü sunamayabilir.

İstatistiksel makine çevirisi sistemlerinin en önemli dezavantajlarından biri eğitim materyali olarak kullanılacak olan derlemlerin tasarım maliyetlerinin oldukça yüksek olmasıdır. Aynı zamanda, derlem tasarımı sürecinde yazılım uzmanları ve dilbilimciler birlikte çalışmalıdır. Bir diğer dezavantaj ise, derlemlerin dil-bağımlı olmasıdır. Başka bir deyişle, makine tarafından öğrenilen ve belli bir dil çifti için geçerli olan dilsel bilgi, diğer diller için geçerli olmayabilir. Artalanda eğitim materyali olarak yer alan derlemler hangi dil çiftlerinde daha fazla ise, o diller arasında yapılan çeviriler daha başarılıdır. Dünya çapında en az 300 milyon kişi tarafından konuşulan İngilizce, Çince ve İspanyolca gibi diller yaygın/majör dil olarak adlandırılır. Yaygın diller aynı zamanda bilim, teknoloji ve ticarete de yön veren dillerdir. Bu diller dilsel kaynak/materyal bakımından da oldukça zengin olup, paralel derlemlerini oluşturmak daha kolaydır. Örneğin, istatistiksel bir makine çevirisi sisteminde İngilizce-İspanyolca dil çifti arasında çeviri yapmak için yeterli büyüklükte derlem mevcuttur, böylece makine çevirisi sistemi gerekli dilsel yapıları rahatlıkla öğrenir ve daha başarılı bir erek metin üretebilir.

Dünya nüfusuyla kıyaslandığında, oransal olarak daha az konuşulan dillere ise azınlık dilleri/minör diller adı verilir. Gürcüce, Yunanca, Bulgarca gibi diller azınlık dillerine örnektir. Bu diller “kaynaktan yoksun diller” olarak da adlandırılır, bir başka deyişle, bu dilleri içeren büyük boyutta paralel derlemler oluşturmak oldukça zordur. Bu nedenle, azınlık dillerini kapsayan istatistiksel makine çevirileri, yaygın diller arasında gerçekleştirilen çeviriler kadar başarılı olamaz. Aynı nedenle, yani yeterli dilsel materyalle öğrenmelerini

gerçekleştiremediklerinden, istatistiksel sistemler, dil çiftlerinden biri minör dil, diğeri ise majör dil olduğunda yine akıcı çeviriler üretmez. Özellikle majör bir dilden minör bir dile çeviri anlamsal düzeyde hatalı çıktılar verebilir. Dolayısıyla, istatistiksel makine çevirisi sistemlerinde düzeltmenlik çabası dil çiftine bağlıdır ve majör diller arasında gerçekleştirilen çevirilerde, düzelti gereksinimi görece düşükken; minör diller söz konusu olduğunda makine çevirisi çıktısı üzerinde daha fazla art düzenleme yapmak gerekebilir.

2. 1. 4. Sinir Ağı Tabanlı Makine Çevirisi Sistemleri

Sinir ağı tabanlı makine çevirisi (İng. *Neural MT*), en gelişmiş makine çevirisi yaklaşımı olarak kabul edilir. 2015 yılından günümüze kadar geçen süreçte sinir ağı tabanlı sistemler önemli gelişmeler kaydetmiştir. Bu sistemlerde, genellikle tam bir tümce şeklinde olan bir sözcük dizisinin kaynak metnin en olası çevirisi olarak belirlenmesi amacıyla büyük bir yapay sinir ağının kullanıldığı bir çeviri yaklaşımı benimsenir. İstatistiksel makine çevirisi sistemlerinin geliştirilmesi için daha fazla bellek kaynağına ve zamana gereksinim vardır. Oysa, sinir ağı tabanlı makine çevirisi, çeviri başarımını en üst düzeye çıkarabilmek için tüm bileşenlerini (sinir ağı katmanları) uçtan uca eğitir. Sinir ağı tabanlı sistemleri eğitmek için, öncelikle bir yapay sinir ağı oluşturulur. Bu ağın bileşenleri, tıpkı bir insanın sinir hücreleri gibi, birbirlerine içsel olarak bağlanmış yazılımsal sinir hücrelerinden oluşur. Sisteme verilen kaynak ve erek metin tümcelerine bağlı olarak, sinir ağı içerisindeki ağırlıklar devamlı olarak güncellenir ve böylece artalandaki tüm eğitim materyali kullanıldığında sinir ağı, çeviriyi gerçekleştirebilir duruma gelir. Sinir ağı tabanlı sistemler, gelişmişlik düzeyleri ile geleneksel makine çevirisi yazılımlarını geride bırakmıştır.

İstatistiksel makine çevirisi sistemi olarak hayatına başlayan *Google Çeviri* de 2016 yılından beri sinir ağı tabanlı bir sistem olarak çalışmaktadır. Alanyazında oldukça yeni bir kavram olan sinir ağı tabanlı makine çevirisi sistemlerinin çalışma mantığını somutlaştırmak için, sıfır atımlı çeviri (İng. *zero-shot translation*) ve aradilden çeviri (İng. *pivot translation, interlingua*) kavramlarını açıklığa kavuşturmak gerekir. Aradil ya da köprü dili (İng. *pivot language*) bir dilden diğeri dile çeviri yapılırken kullanılan yapay ya da doğal dildir. Makine çevirisi sistemlerinde kaynak ve erek dilde makine öğrenmesini sağlayacak kadar dilsel veri yoksa aradil kullanımıyla çeviri gerçekleştirilebilir. Örneğin, kaynak dil Bulgarcadan erek dil Gürcüceye çeviri yaptığımızı düşünelim. Her iki dil de minör dil statüsünde olduğundan ve sistem artalanında makine öğrenmesini sağlayabilecek büyüklükte paralel derlem bulunmadığından, “İngilizce” aradil olarak kullanılarak makine çevirisi gerçekleştirilebilir. Bu durumda, makine çevirisi sistemi tarafından Bulgarca kaynak metin önce İngilizceye çevrilir; ardından üretilen İngilizce metin ise erek dil Gürcüceye çevrilir. Aradilden çeviri, geleneksel çeviri süreçlerinde olduğu gibi makine çevirisi sistemlerinde de bilgi kayıplarına, hata ve belirsizliklere neden olabilir.

Sinir ağı tabanlı sistemlerin gelişimiyle birlikte sıfır atımlı çeviri kavramı ortaya atılmıştır. Kaynaktan yoksun dilleri içeren çeviri süreçlerinde, istatistiksel makine çevirisi sistemleri gibi sinir ağı tabanlı makine çevirisi sistemlerinin de zorlandığı görülmüştür. Çok dilli sinir ağı tabanlı makine çevirisi yaklaşımları son yıllarda geliştirilmeye başlanmış ve sıfır

atımlı çeviri yaklaşımı gündeme gelmiştir. Sıfır atımlı çeviri (Johnson, vd., 2017, ss. 339-351), sistemin, eğitim materyalleri içerisinde hiç karşılaşmadığı dil kombinasyonları arasında çeviri yapabildiğini sağlayan bir modeldir. Örneğin, Türkçeden Bulgarcaya yapılacak bir çeviride, Türkçe – Bulgarca dil çiftinde yeterince paralel metin yoksa bu durum sistem başarımını olumsuz yönde etkilemez, çünkü birçok başka dilden öğrenme materyali sistem içinde mevcuttur ve dilden bağımsız olarak çok sayıda dil çiftinde bu materyaller kullanılır.

Sıfır atımlı çeviri yaklaşımı birçok çalışmada test edilmiştir. İtalyanca-Rumence dil çiftinde sıfır atımlı yöntemin başarımını ölçmek için yapılan bir çalışmada Rumence, İtalyanca, Almanca, Hollandaca ve İngilizceyi içeren çok dilli bir veri tabanı oluşturulmuştur. Çalışmaya dahil edilen tüm dillerin İngilizce ile paralel derlemleri bulunmaktadır. Deneysel çalışmalar sonucunda, çok dilli sinir ağı tabanlı makine çevirisi sisteminde kullanılan tekrarlamalı yöntemin, çeviri yönü fark etmeksizin sıfır atımlı çeviride daha kaliteli sonuçlar verdiği görülmüş; bu yöntemin aradilden çeviri kullanılarak yapılan makine çevirisi sistemlerinden de başarılı olduğu gözlemlenmiştir (Lakew, vd., 2018). Sıfır atımlı çeviri yöntemini kullanan sinir ağı tabanlı makine çevirisi sistemleri görece daha başarılı ve erek odaklı çıktılar üretebildiğinden art düzenleme çabası da diğer sistemlere göre daha düşüktür. Bu nedenle, bu sistemler, özellikle hesaplamalı dilbilim alanında bir paradigma değişimine neden olmuştur. Çevirmenlerin de makine çevirisi sistemlerine güveni artmış, bu sistemler çevirmenler tarafından yoğun bir şekilde kullanılmaya başlanmıştır.

Günümüzde en sık kullanılan iki makine çevirisi sistemi olan *Google Çeviri* ve *Yandex Çeviri* sinir ağı tabanlı yapılar kullanmaktadır. Ancak, *Yandex Çeviri* sinir ağı tabanlı makine çevirisi yanında, istatistiksel makine çevirisi sistemlerinden de yararlanarak karma (hibrit) bir modelle çalışmaktadır. “İki model bir modelden daha iyidir” sloganıyla hareket eden *Yandex Çeviride*, kullanıcı çevrilecek metni platforma girer girmez, metin hem istatistik hem de sinir ağı tabanlı sistemlere gönderilir. Her iki sistemden elde edilen sonuçlar, *CatBoost*⁵ adlı makine öğrenmesi yöntemine dayalı bir algoritma ile değerlendirilir. Algoritma, tümce uzunluğundan sözdizimine kadar bir dizi unsuru analiz eder. Örneğin, kısa ifadeler ve nadir kullanılan sözcükler istatistiksel model tarafından daha iyi çevrilir. İki çeviri tüm bileşenler açısından karşılaştırılır ve kullanıcıya en iyi olanı gösterilir.

Sanılanın aksine, makine çevirisi süreci tam otomatik bir süreç değildir. Gerek kaynak gerekse de erek metin üzerinde düzeltmeler gerçekleştirecek olan uzmanın, kullanılan makine çevirisi sisteminin yeterli ve zayıf yönlerini bilmesi gerekmektedir. Bu bağlamda, farklı makine çevirisi sistemlerinin özelliklerinin bilinmesi, makine çevirisinin tarihsel gelişimine yönelik bilgi sahibi olunması ve makinenin “nasıl düşündüğünün” kavranması daha sağlıklı bir düzenleme sürecini de beraberinde getirecektir. Çevirmenlerin makine çevirisi sistemlerinin çalışma mantığını anlaması, başarılı erek metinlerin oluşturulması için önem taşımaktadır. Ancak, sinir ağı tabanlı sistemler her ne kadar teknik alan için gerek bağlamsal gerekse de terminoloji düzeyinde başarımlar sağlayabilse de çevirmenlerin makine çevirisini yalnızca

⁵ <https://yandex.com/dev/catboost/> (Erişim tarihi: 20.05.2023)

“çevirmene yardımcı araç” olarak görmesi ve makine çevirisi çıktılarını sorgulayarak gerekli düzelteleri yapması oldukça önemlidir.

Günümüzde makine çevirisi üzerine çalışan araştırmacılar aynı zamanda makine çevirisinin genel kullanıcı tarafından da verimli biçimde kullanılabilmesi için çaba sarfetmektedir. Örneğin, Ottawa Üniversitesi çeviribilim bölümünde görevli olan Lynne Bowker, “Makine Çevirisi Okur Yazarlığı Projesi” (İng. *Machine Translation Literacy Project*)’ni hayata geçirmiştir. Bowker, projenin amacını “çevirmen olmayan kişilerin bilinçli birer makine çevirisi sistemi kullanıcısı olmasına yardımcı olmak” olarak açıklamaktadır. Avrupa Birliği tarafından hayata geçirilen bir proje olan *MultiTraiNMT*’nin amacı da AB bünyesindeki çok dilli vatandaşların daha bilinçli makine çevirisi sistemi kullanıcıları olmalarına destek olmaktır (Bowker, 2022).

Makine çevirisi sistemleri alan bazlı ve kullanıcı odaklı olarak da tasarlanabilmektedir. Kullanıcı beklentilerinin ön plana alındığı yapıda tasarlanan sistemlerin temel amacı kullanıcıların çeviri deneyimini iyileştirmektir. Bu sistemlere “özel amaçlı sistemler” ve “konuşma çevirisi sistemleri” örnek olarak verilebilir (Ping, 2011, ss, 164-165). Özel amaçlı sistemler, belli bir uzmanlık alanında kullanım için tasarlanmıştır. Örneğin, Kanada’da yayınlanan meteoroloji bültenlerinin İngilizce-Fransızca dil çiftinde çevirileri için hazırlanmış bir sistem olan *METEO*, özel amaçlı bir makine çevirisi yazılımıdır. Konuşma tanıma, konuşma üretme ve makine çevirisi teknolojilerini bir araya getiren ve kaynak dilde konuşmayı doğrudan sesli olarak erek dile aktaran *JANUS* ise konuşma çevirisi yapan makine çevirisi sistemlerine örnek olarak verilebilir.

Doğal dil işleme yöntemlerinin gelişimi ve buna bağlı olarak sinir ağı tabanlı yaklaşımların yaygınlaşması sonucunda, birçok makine çevirisi sisteminin tasarımında ilerlemeler kaydedilmiş; yeni sistemler hayata geçmiştir. 500 milyonu aşkın kullanıcı sayısı⁶ ile gerek genel gerekse de uzman kullanıcılar tarafından günümüzde en fazla tercih edilen makine çevirisi yazılımı *Google Çeviri* olsa da özellikle uzman kullanıcılar görece yeni bir sistem olan *DeepL Translator*’a yönelmiştir. *DeepL Translator*’ı yaygın kullanım bulmuş diğer makine çevirisi yazılımlarından ayıran bazı özellikler bulunmaktadır. Bu özelliklere değinilerek, doğal dil işleme ve sinir ağı tabanlı modellerin makine çevirisi sistemlerinin tasarımında geldikleri son noktanın incelenmesi önem taşımaktadır.

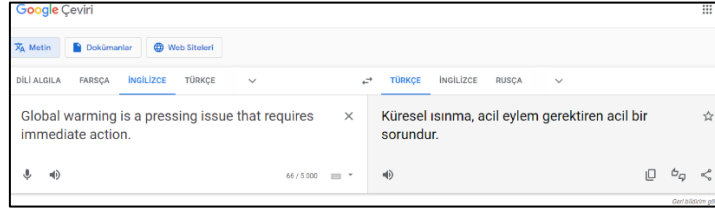
DeepL Translator, kaliteli çeviriler gerçekleştirmek için ileri düzey makine öğrenmesi yöntemlerini ve sinir ağlarını kullanan çevrimiçi bir çeviri hizmetidir. Platform, 2017 yılında Alman teknoloji şirketi *DeepL GmbH* tarafından hayata geçirilmiştir. *DeepL Translator*, ilk olarak yedi Avrupa Birliği dilinde (İngilizce, Almanca, Fransızca, İspanyolca, İtalyanca, Hollandaca ve Lehçe) çeviri yapmaya başlamış, ancak şu anda Türkçeyi de içine alan 31 dilde çeviri yapabilmektedir⁷. *DeepL Translator* çevirmene sunduğu özellikler açısından diğer makine çevirisi yazılımlarından ayrışmaktadır. Burada kastedilen ayrışma, makine çevirisi

⁶ <https://www.k-international.com/blog/google-translate-facts/> (Erişim tarihi: 20.05.2023)

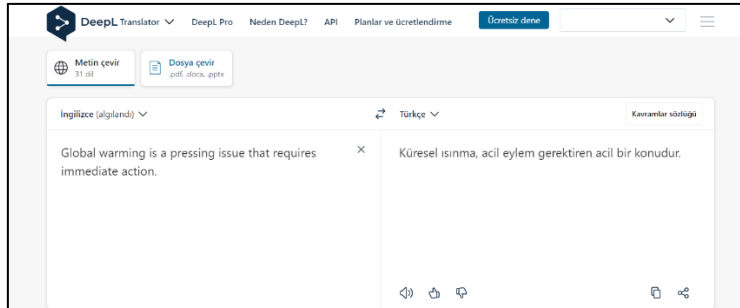
⁷ <https://www.deepl.com/translator> (Erişim tarihi: 20.05.2023)

yazılımlarının sunduğu çıktının başarımı ile ilgili değildir; çünkü makine çevirisi platformlarının başarımını etkileyen bir dizi etmen bulunmaktadır. Bunlardan bazıları metin türü, konu alanı, dil çiftleri, kaynak metnin karmaşıklık düzeyi ve sistemlerin geliştirilmesinde kullanılan makine öğrenmesi yöntemleridir. Örneğin, bir makine çevirisi yazılımı tıp metinlerinde daha iyi çıktılar üretebilirken; başka bir makine çevirisi hizmeti haber metinleri çevirisinde daha başarılı olabilir.

"En iyi" makine çevirisi sistemi diye bir şey yoktur. Başarım, kaynak metnin makine çevirisi modellerini eğitmede kullanılan verilere ne kadar benzediğine ve algoritmalara bağlıdır (Intento, 2021). Makine çevirisi sistemlerinin başarım ölçümleri bir dizi metrikle gerçekleştirilmektedir. Bunlardan bazıları, anlamsal benzerliğin değerlendirildiği *BERTScore*, *TER*, *PRISM* ve *COMET*'tir (a.g.e., 18). Dolayısıyla, mevcut makine çevirisi hizmetlerinin çeviri başarılarının ölçülmesi tek yönlü analizlerle mümkün olmamaktadır. Ancak, ilgili platformların kullanıcıya sunduğu özellikler açısından karşılaştırılması yapılabilir. Görsel 1 ve 2'de aynı kaynak metnin *Google Translate* ve *DeepL Translator* üzerinden alınmış çıktıları görülmektedir:

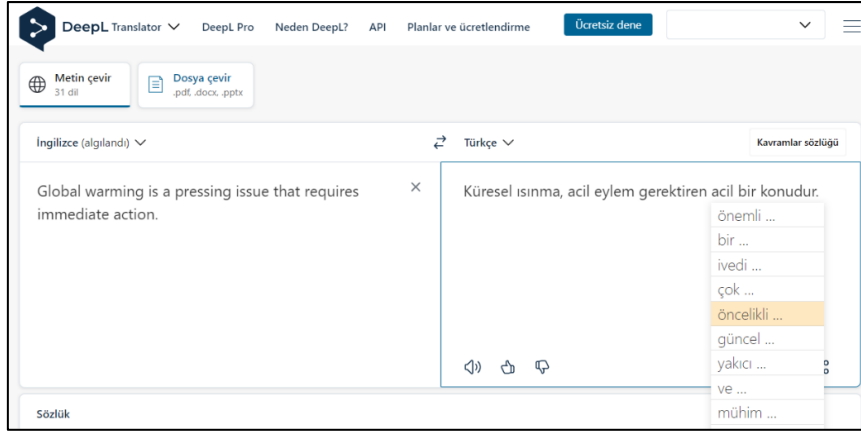


Görsel 1. İngilizce-Türkçe dil çiftinde makine çevirisi örneği (*Google Çeviri*).



Görsel 2. İngilizce-Türkçe dil çiftinde makine çevirisi örneği (*DeepL Translator*).

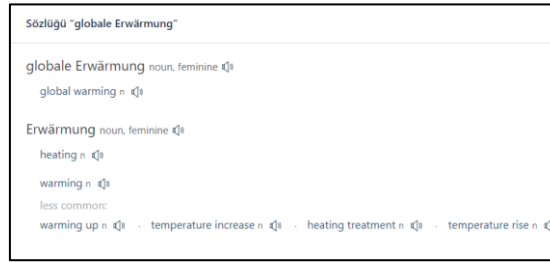
Görsel 1 ve 2'de yer alan ve iki makine çevirisi platformundan alınan erek metinlerde belirgin bir fark göze çarpmamaktadır. Ancak, *DeepL Translator* kullanıcıya sözcük ve tümce bazlı farklı çeviri seçenekleri de sunmaktadır. Görsel 3'te bu durum örneklenmiştir:



Görsel 3. İngilizce-Türkçe dil çiftinde makine çevirisi örneği (Eş anlamlı sözcük gösterimi – DeepL Translator)

Görsel 3’te görülebildiği gibi, kullanıcı erek metinde yer alan “acil” sözcüğü yerine program tarafından sunulan başka bir alternatifi de tercih edebilir. Kullanıcı seçimlerine göre erek metin tekrar düzenlenir ve yapılandırılır. Söz konusu bu özellik daha erek odaklı çevirilerin üretimine katkı sağlayabilmektedir.

DeepL Translator’ı diğer makine çevirisi sistemlerinden ayıran bir diğer özellik ise bazı dil çiftlerinde oluşturulmuş terimcelerin kullanıcıya sunulmasıdır. Görsel 4’te İngilizce-Almanca dil çiftinde bir terim sorgusu yer almaktadır:



Görsel 4. İngilizce-Almanca dil çiftinde terimce örneği (DeepL Translator).

Terimcelere erişim özelliği de çeviri sürecini daha verimli kılabilir. Aynı zamanda, kullanıcı sistem üzerinde kendi kavramlar sözlüğünü de oluşturabilmektedir. Halihazırda terimce özelliği Türkçe için desteklenmemektedir.

DeepL ayrıca kullanıcıya bir yazma asistanı da sunmaktadır. Henüz deneme sürümü aşamasında olan DeepL Write⁸, dilbilgisi ve yazım hatalarını düzeltme, tümceleri yeniden yazma, uygun dil kesitini seçme, daha yaratıcı dilsel seçenekler sunma gibi birçok özelliğe sahiptir. Bu modül de kullanıcıya makine çevirisi çıktısının düzenlenmesinde yardımcı olmaktadır. Halihazırda Almanca ve İngilizce dillerinde kullanılabilen özelliğin daha fazla dili kapsamaya planlanmaktadır.

Doğal dil işleme teknikleri ışığında sinir ağı tabanlı yöntemlerin gelişmesi ile makine çevirisi sistemleri gün geçtikçe daha kapsamlı hale gelmekte ve kullanıcıya daha fazla özellik

⁸ <https://www.deepl.com/write> (Erişim tarihi: 20.05.2023)

sunabilmektedir. Çevirmenler tarafından kullanılan doğal dil işleme tabanlı bir diğer araç ise otomatik konuşma tanıma ve sözlü çeviri sistemleridir.

3. Otomatik Konuşma Tanıma Sistemleri ve Makine Destekli Sözlü Çeviri

Otomatik konuşma tanıma, konuşulan dili insan müdahalesi olmadan yazılı metne veya bilgisayar tarafından okunabilir biçime dönüştürmek için makine öğrenimi algoritmalarını kullanan bir teknolojidir. Otomatik konuşma tanıma aynı zamanda doğal dil işlemede kullanılan ve bilgisayarların konuşma dilini anlamasını ve bu dile ait verilerin bilgisayarlar tarafından işlenmesini sağlayan bir yöntemdir. Bu yöntem makine destekli sözlü çevirinin de ilk adımını oluşturur.

Claudio Fantinuoli (2023) makine destekli sözlü çeviri sürecini şöyle açıklar:

Sözlü çeviri sürecinde basamaklı bir yaklaşım söz konusudur; süreç bilgisayar programları ile modellenen alt parçalara ayrılır. Bu süreçte üç ayrı bileşenden söz edilebilir: Otomatik konuşma tanıma, söylenen sözcükleri yazılı metne çevirir, makine çevirisi üretilen bu yazılı metni kaynak dilden erek dile aktarır ve metinden konuşma üretimi (İng. *text-to-speech synthesis*) yöntemiyle çevirinin sözlü bir versiyonunu üretir. Bu bileşenler, bir işlemin çıktısının bir sonrakinin girdisi durumunda olduğu “basamaklı çeviri” adı verilen bir süreçte ardışık olarak uygulanır.

Fantinuoli, son zamanlarda makine destekli sözlü çeviride önemli iyileştirmelerin gözlemlendiğini; ancak gerçek zamanlı çok dilli iletişimin getirdiği zorluklar nedeniyle çeviri kalitesinin istenen düzeye ulaşamadığını belirtir (a.g.e.)⁹. Makine destekli sözlü çeviride kullanılan diğer bir yöntem ise uçtan uca sözlü çeviridir. Bu yöntemde makine öğrenmesi algoritmaları kaynak ve erek dildeki ses dosyaları ile eğitilir. Eğitimin sonucunda ortaya çıkan model, kendisine girdi olarak verilen bir ses dosyasını ara bir makine çevirisi sistemine gerek duymadan erek dildeki ses dosyasına dönüştürür. Halihazırda emekleme döneminde olduğu söylenebilecek bu yöntem, basamaklı yaklaşıma kıyasla çok daha hızlı sonuçlar üretse de uçtan uca yöntemde çıktı kalitesi daha düşüktür.

Yapay zekâ ve makine çevirisi destekli sözlü çeviri 2020 yılında düzenlenen Tokyo olimpiyatlarında kullanılmıştır. Japon teknoloji devi Panasonic, *100BANCH* adlı bir firma ile iş birliği yaparak *Fukidashi* adlı prototip bir yapay zekâ çevirmeni üretmiştir. iPad benzeri bu cihaz, kullanıcılar için eş zamanlı çeviri hizmeti sağlamıştır. Cihaz, sporcular, personel ve Tokyo'ya gelen ziyaretçilerin dil bariyerini aşmasına yardımcı olmuştur. Benzer şekilde, 2022 Pekin Kış Olimpiyatlarında *iFLYTEK* adlı Çinli çeviri şirketi, olimpiyat oyunlarının resmi çevirmeni olarak belirlenmiştir. *iFLYTEK*, beş dilde otomatik sözlü çeviri hizmeti sunan Xiaoqing'i piyasaya sürmüştür (Language Connections, 2022). Otomatik konuşma çevirisi, uluslararası konferanslar, iş toplantıları ve seyahatler gibi çeşitli ortamlarda farklı diller konuşanlar arasındaki iletişimi kolaylaştırma potansiyeline sahiptir. Doğal dil işleme alanında

⁹ Claudio Fantinuoli, makine destekli sözlü çeviriye yardımcı bir araç geliştirmiştir. Araca, <https://www.machine-interpreting.com/> adresinden erişilebilmektedir (Erişim tarihi: 20.05.2023).

gerçekleşen gelişmeler gelecekte makine destekli sözlü çeviri çalışmalarının artmasını sağlayabilecektir.

4. Otomatik Metin Özetleme

Otomatik metin özetleme, özetleyerek çeviri kavramı ile bağlantılı olarak ele alınmalıdır. Özetleyerek çeviri (İng. *summary translation, précis translation*), en temel anlamıyla çevirmenin daha az sözcük kullanarak kaynak metnin özünü aktarmaya çalıştığı, daha uzun bir metnin veya belgenin kısaltılmış ya da özet çevirisidir. Özet çevirinin genel amacı, gereksiz ayrıntılara ve bilgilere yer vermeden kaynak metnin temel fikirlerinin açık ve özlü bir şekilde sunulmasıdır. Özetleyerek çeviri, genellikle içeriğe hızlı bir genel bakışın gerekli olduğu araştırma makaleleri, yasal belgeler veya teknik raporların erek dile aktarılmasında kullanılır. Çevirmenler, teknik alanda yazılmış olan, terim yoğunluklu bir kaynak metni anlamak için de özetleyerek çeviriye başvurur. Ayrıca, çeviri eğitiminde ve dil öğreniminde tercih edilen bir yöntemdir.

Çeviribilim alanyazınında konuyla ilgili görece yetersiz sayıda çalışma olsa da (Shreve, 2006) özetleyerek çevirinin doğası ve uygulanabilirliği ile ilgili çalışmalar mevcuttur (Bowker ve McBride, 2017; Vandermeulen, 2020). Çevirmenlerin sürekli olarak zaman baskısı altında çalışmak zorunda kalması ve profesyonel iş yaşamının gereklerine hızlıca ayak uydurmak durumunda olması dolayısıyla, çeviriye hızlı ve doğru şekilde gerçekleştirme konusunda da eğitim almaları gerekmektedir. Bu bağlamda çevirmenlerin “özetleyerek yazma” edincini kazanmaları oldukça önemlidir (Bowker ve McBride, 2017). İspanyolca - Hollandaca dil çiftinde aday çevirmenlerle yapılan bir anket çalışmasında özetleyerek çevirinin zorlukları şu şekilde belirlenmiştir: (1) kaynak metnin anlaşılması; (2) erek metinde yer alması gereken kaynak metin unsurlarının seçimi; (3) erek metnin oluşturulma süreci. Sonuç olarak, çevirmen adayları özetleyerek çeviri etkinliğinin kendilerini zorladığını dile getirmiştir (Vandermeulen, 2020, ss. 16-18). Dolayısıyla, özetleyerek çeviri çevirmenin yalnızca bireysel çabasına bırakıldığında normal bir çeviri sürecine kıyasla daha zorlayıcı olabilmektedir.

Andrew Chesterman’a göre (2016, s. 152) özetleyerek çeviri, hesap verebilirlik konusuyla da ilgilidir. Bu işlem, çevirmenin çevrilmesi gerekenleri seçip uygun şekilde yapılandırma sorumluluğunu içerdiğinden, öznel kararlar almasına da neden olabilir. Özetleyerek çeviri, doğal olarak "sıradan" bir çeviriye kıyasla çevirmenin "kendi zihnindeki" kaynak metne daha benzer bir metin üretmesi anlamına gelebilir. Bu da çeviride manipülasyonu beraberinde getirebilir.

Sonuç olarak, gerek özetleyerek çeviri işleminin (özellikle teknik bir alanda özet çeviriye dahil edilmesi gereken kilit unsurların seçimi konusunda) çevirmeni zorlaması, gerekse de çevirmen çevirisini yaptığı alanda uzman olmadığından kaynak metin unsurları seçiminde yeterince nesnel davranamaması nedeniyle otomatik metin özetleme sistemleri bu konuda oldukça işlevsel şekilde çevirmene destek olmaktadır. Söz konusu sistemler kaynak metnin özetlenmesi amacıyla çevirmenler tarafından kullanılabilir. Kaynak metin özetinin erek dile çevrilmesiyle özetleyerek çeviri gerçekleştirilmiş olur.

Otomatik metin özetleme, birçok açıdan çevirmene yardımcı olabilir:

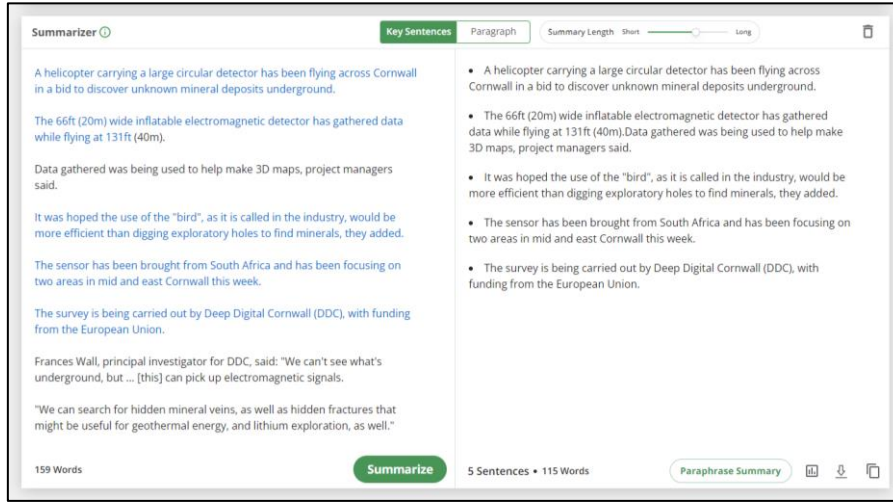
- **Kaynak metin hakkında genel bir anlayış oluşturulması:** Çeviriye başlamadan önce çevirmen, üzerinde çalışacağı metnin neyle ilgili olduğu konusunda fikir elde etmek için bir otomatik metin özetleme aracı kullanabilir. Bu araç, çevirmenin metnin temel fikirlerini anlamasına ve anahtar kavramlarla özel alan terminolojisini belirlemesine yardımcı olabilir.
- **Erek metne doğru ve eksiksiz şekilde aktarılması gereken önemli noktaların belirlenmesi:** Otomatik metin özetleme, çevirmenin kaynak metindeki önemli bilgileri belirlemesine yardımcı olabilir. Bu işlem, özellikle kilit bilgileri belirlemenin zor olabileceği uzun teknik ve bilimsel belgelerle çalışırken faydalı olabilmektedir.
- **Özet çevirinin oluşturulması:** Otomatik metin özetleme, çevirmenin özet çeviri üretmesine yardımcı olabilir. Bu özet çeviri, çeviri sürecinde bir başlangıç noktası olarak yer alabilir; çevirmenin daha özlü ve anlaşılır bir çeviri oluşturmasına destek olabilir.
- **Erek metnin okunabilirliğinin artırılması:** Otomatik metin özetleme, erek metnin daha erek odaklı, okunabilir ve işlevsel olmasına katkı sağlayabilir. Çevirmen, tekrarlı, gereksiz veya ilgisiz bilgileri belirleyip kaldırarak okunması ve anlaşılması daha kolay bir erek metin oluşturabilir.

Çeviri sürecini desteklemenin yanı sıra, otomatik metin özetleme “özetleyerek çeviri” gerektiren çeviri işlerinde de doğrudan çözüm sunabilir. Örneğin, basın alanında yapılan çeviriler ve uzun işletme raporlarının çevrilmesi süreçlerinde özetleyerek çeviri işlemine başvurulması istenebilir. Bu noktada otomatik metin özetleme araçları çevirmene destek olur. Çevirmen kaynak metin özetini alıp çevirir. Bu da çevirmene zaman kazandırır. Dolayısıyla, otomatik özetleme ve özetleyerek çeviri oldukça bağlantılı iki işlemdir.

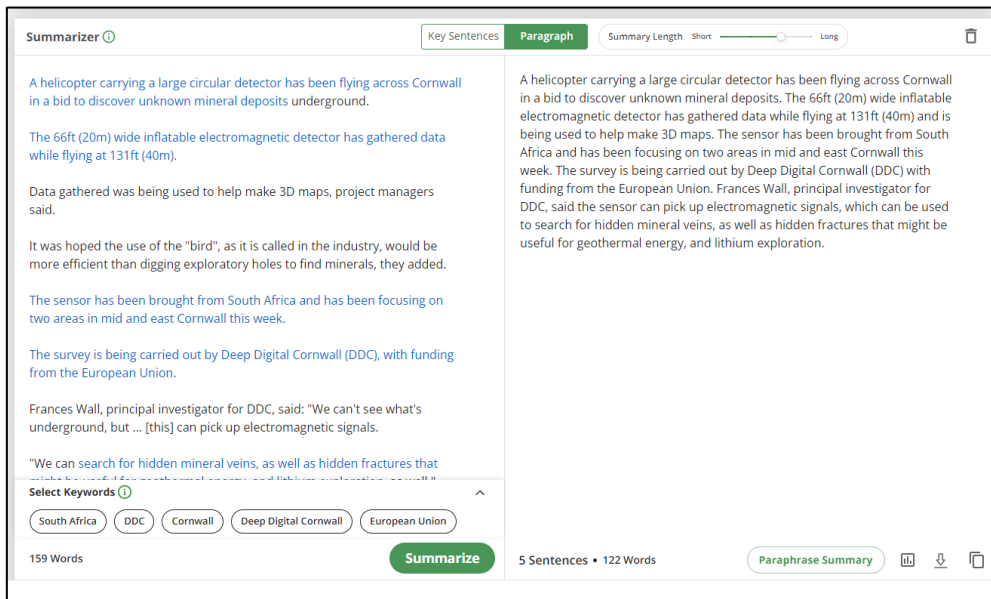
Otomatik metin özetleme iki şekilde gerçekleştirilebilir. Çıkarımsal özetleme (İng. *extractive summarization*) en önemli tümceler ve ifadelerin özgün metinden seçilip çıkarılmasını ve ardından bir özet oluşturmak için bu unsurların birleştirilmesini içerir. Tümceler genellikle konuyla ilgileri, kullanım sıklıkları ve önem dereceleri gibi belirli ölçütlere göre seçilir. Soyutlayıcı özetleme (İng. *abstractive summarization*) yönteminde ise, özgün metinde bulunmayan yeni tümce ve ifadeler içeren bir özet oluşturulmaya çalışılır. Sistem, özgün metnin anlamını kavramaya çalışır ve ardından farklı sözcükler ve tümce yapılarını kullanarak aynı anlamı içeren bir özet oluşturur. Soyutlayıcı özetleme, metnin daha derinlemesine anlaşılması ile akıcı ve tutarlı bir dil üretme becerisini gerektirdiği için genellikle çıkarımsal özetlemeden daha zorlayıcı bir doğal dil işleme süreci olarak kabul edilir.

Çıkarımsal özetlemede özet, özgün metnin bir alt kümesidir; soyutlayıcı özetlemede ise özet, özgün metnin açıklanmasıyla ya da anlamsal bir temsile kodlanmasıyla yeni baştan üretilir (Eisenstein, 2018, ss. 393-394). Çeviri odaklı düşündüğümüzde ise soyutlayıcı özetlemenin çıkarımsal özetlemeye göre çevirmene daha faydalı olabileceğini söyleyebiliriz; çünkü soyutlayıcı özetlemede “yeniden yazma” (İng. *paraphrasing*) temel alındığından, daha özlü, anlaşılır, doğal ve sadeleştirilmiş bir özet üretilebilmektedir. Çıkarımsal özetlemede ise bir dizi ölçüt dikkate alınarak metnin ana hatları belirlenmekte ve seçilen bu temel bilgiler özeti

oluşturmaktadır. Günümüzde sık kullanılan otomatik özetleme araçları arasında *QuillBot*¹⁰, *Text Compactor*¹¹ ve *TLDRthis*¹² sayılabilir.



Görsel 5. Çıkarımsal özetleme örneği (özgün metinden önemli tümcelerin doğrudan alınması – *QuillBot* özetleme aracı üzerinden)¹³.



Görsel 6. Soyutlayarak özetleme örneği (özgün metni kısaltarak yeniden yazma – *QuillBot* özetleme aracı üzerinden).

Özetleme araçları kullanıcıya oluşturulacak özetin uzunluğunu belirleme şansı da tanımaktadır. Kullanıcı aynı zamanda aracı kullanarak anahtar sözcükleri ya da alana özgü

¹⁰ <https://quillbot.com/summarize> (Erişim tarihi: 20.05.2023)

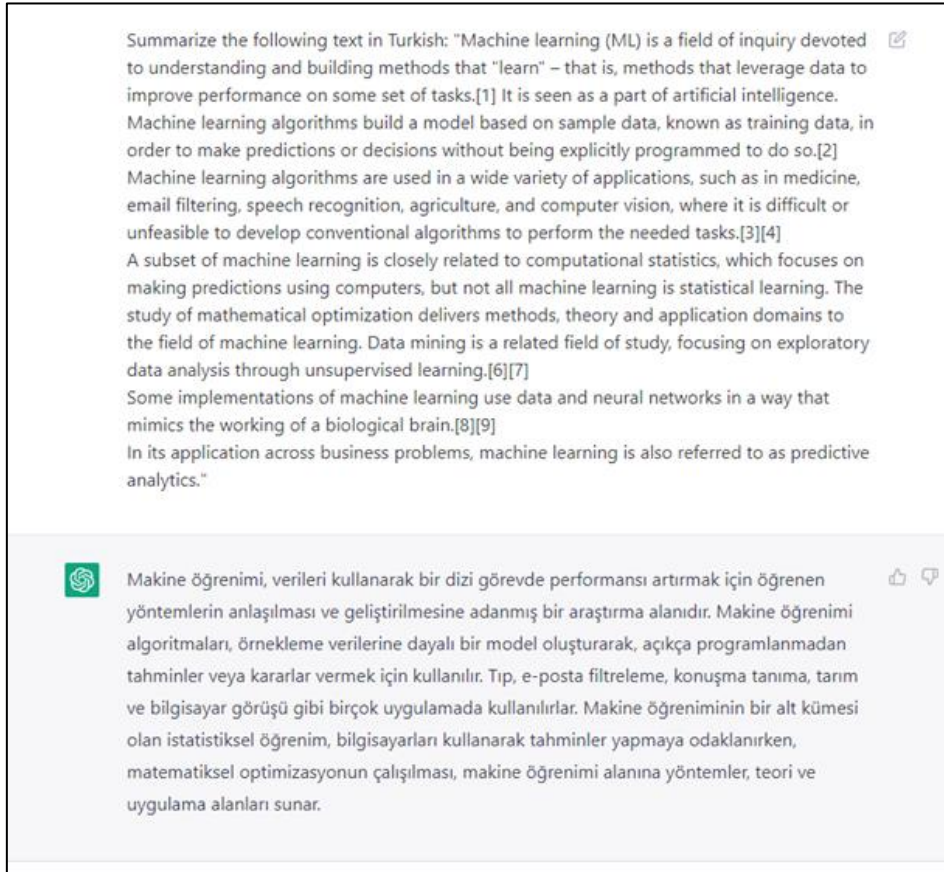
¹¹ <https://www.textcompactor.com/> (Erişim tarihi: 20.05.2023)

¹² <https://tldrthis.com/> (Erişim tarihi: 20.05.2023)

¹³ Özgün metin, <https://www.bbc.com/news/uk-england-cornwall-64810962> adresinden alınmıştır (Erişim tarihi: 20.05.2023).

terminolojiyi belirleyebilir; özeti farklı bir dilsel içerikle yazdırabilir. Dolayısıyla, otomatik özetleme araçları günümüzde çevirmenin iş istasyonunda yerini almıştır.

Günümüzde doğal dil işleme tekniklerinin ulaştığı son nokta olarak *ChatGPT* örnek verilebilir. 2022 yılı Kasım ayında San Francisco merkezli teknoloji şirketi *OpenAI* tarafından tanıtılan *ChatGPT*, yapay zekâ tabanlı bir sohbet robotu olup, kullanıcıların sorularını yanıtlamaktadır¹⁴. *ChatGPT* aynı zamanda hatalarını kabul etmekte, uygunsuz sorguları yanıtlamayı reddetmekte ve hatta yalan bile söyleyebilmektedir. Kullanıcıların sorgularını anlayabilmek ve gerekli yanıtları verebilmek için yapay zekâ ve doğal dil işleme tekniklerini kullanan *ChatGPT*, insan elinden çıkmış kadar doğal metinleri birçok dilde üretebilmektedir. Araç aynı zamanda çevirmene yardımcı olabilecek bir dizi doğal dil işleme bileşeni sunmaktadır. Bunlardan biri de araca verilen kaynak dildeki metnin doğrudan erek dilde özetlenmesidir. Görsel 7’de görüldüğü gibi, “Summarize the following text in Turkish.” (Tr. Aşağıdaki metni Türkçe özetle.) şeklindeki basit bir komutla bu işlemi gerçekleştirebilmektedir.



Görsel 7. *ChatGPT*^{15,16} üzerinde özetleyerek çeviri işlemi.

¹⁴ <https://openai.com/blog/chatgpt> (Erişim tarihi: 20.05.2023)

¹⁵ <https://chat.openai.com/chat> (Erişim tarihi: 20.05.2023)

¹⁶ Özgün metne, https://en.wikipedia.org/wiki/Machine_learning adresi üzerinden erişilebilmektedir (Erişim tarihi: 20.05.2023).

Sonuç

Bu çalışmada doğal dil işleme ve çeviri etkinliği arasındaki ilişki yapay zekâ alanındaki güncel gelişmeler ışığında ele alınmaya çalışılmıştır. Doğal dil işleme teknikleri, dillerarası çeviri ile ilgili süreçleri önemli ölçüde etkilemiş, makine öğrenmesi algoritmaları, istatistiksel ve hesaplamalı modeller gün geçtikçe daha işlevsel erek metinlerin üretimine olanak sağlamaya başlamıştır.

Kuşkusuz, doğal dil işlemenin çevirmen açısından görünür olduğu en önemli alan makine çevirisidir. Günümüzde birçok makine çevirisi sistemi, sinir ağı tabanlı modeller kullanmaktadır. 2015 yılından itibaren yoğun bir şekilde kullanılmaya başlanan sinir ağı tabanlı makine çevirisi, geçmişteki makine çevirisi yaklaşımlarına (kural, örnek ve istatistik tabanlı sistemler) göre çok daha başarılı sonuçlar vermektedir. Kaynak ve erek dillerde paralel metinlerden oluşan büyük veri kümeleri ile eğitilen sinir ağı tabanlı sistemler, diller arasında nasıl çeviri yapılacağını öğrenmek için yapay sinir ağlarını kullanır.

Sıfır atımlı çeviri yöntemi ile çalışan bu sistemler, seçilen dil çiftinde sistem eğitiminde kullanılan az miktarda veri olsa da, diğer dil çiftlerinin eğitiminde kullanılan verilerden faydalanarak başarılı sonuçlar üretir. Dolayısıyla, sinir ağı tabanlı makine çevirisi, kaynak bakımından yoksun diller söz konusu olduğunda da diğer makine çevirisi yaklaşımlarına göre daha başarılıdır. Ancak, çevirmen-bilgisayar etkileşiminin kesişim kümesinde yer alan makine çevirisinin tam otomatik bir süreç olduğunu düşünmek yanlıştır. Hem çeviri öncesi hem de sonrasında gerekli düzelteleri yapma sorumluluğu insan çevirmene aittir. Bu bağlamda, sistemin daha başarılı bir çıktı üretebilmesi için kaynak metin üzerinde ön düzenlemelerin gerçekleştirilmesi; başka bir deyişle, kaynak metnin dilsel açıdan makine çevirisi sisteminin gereklerine uygun hale getirilmesi önemli bir adımdır. Ayrıca, ham makine çevirisi çıktısının üzerinde sözcük, sözdizimi ve anlam düzeylerinde düzeltelerin yapılması gerekebilmektedir. Bu nedenle çevirmenin kullandığı makine çevirisi sisteminin güçlü ve zayıf yönlerini bilmesi önem taşımaktadır.

Son yıllarda makine destekli sözlü çeviri alanında da dikkat çeken gelişmeler olmuştur. Sözlü çeviri sürecinde çevirmene yardımcı olabilecek araçlar tasarlanmaya başlanmış ve çevirmeni sürece hazırlayabilecek terimce oluşturma araçları gibi birtakım uygulamalar hayata geçirilmiştir. Ancak, makinelerin eş zamanlı çeviri yapabilme kapasiteleri halen sınırlıdır. Yapay zekâ ve doğal dil işleme alanındaki gelişmelere paralel olarak bu konuda da önemli ilerlemeler kaydedilmektedir.

Otomatik özetleme araçları da çevirmen tarafından kullanılacak doğal dil işleme uygulamalarıdır. Bu araçlar özellikle çevirmenin kaynak metni anlamasına ve anahtar terimleri çıkarmasına yardımcı olur. Günümüzde yalnızca metin özetleme değil özetleyerek çeviri yapabilen otomatik araçlar da mevcuttur. Çevirmenin iş istasyonunda yer alan doğal dil işleme tabanlı araçlar bunlarla sınırlı değildir; terim yönetim sistemlerinden yeniden yazma araçlarına kadar birçok uygulamada doğal işleme teknikleri kullanılmaktadır.

Artan çeviri ihtiyacıyla birlikte çeviri süreçlerinin kısalması gerekmektedir ve buna bağlı olarak da çevirmenin doğal dil işleme yöntemlerini kullanan araçlardan yararlanması neredeyse

bir zorunluluk haline gelmiştir. Dolayısıyla, çevirmenin doğal dil işleme kavramına ilişkin farkındalık geliştirmesi ve ilgili uygulamaların işleyiş şekillerini öğrenmesi, söz konusu sistemlerin kısıtları ve kapasitelerini doğru bir şekilde kavramasını sağlayabilecektir.

İleriki çalışmalarda çevirmen odaklı özelleşmiş doğal dil işleme teknikleri, doğal dil işlemenin çeviri kalitesi üzerindeki etkisi ve çeviri piyasasında doğal dil işlemenin konumu gibi konuları merkeze alan incelemeler yapılabilir. Bunun yanında, doğal dil işleme tabanlı sohbet robotları ve makine çevirisi sistemlerinin çeviri kalitesi açısından karşılaştırılmasını kapsayan görgül çalışmalar gerçekleştirilebilir. Ayrıca, doğal dil işleme araç ve yöntemlerinin çeviri bölümlerine ait ders planlarında yer alması, bir yandan çeviribilim öğrencilerinin teknoloji edincini geliştirecek, diğer yandan ise onları günümüz çeviri piyasasına hazırlayabilecektir. Bu nedenle, doğal dil işleme tabanlı araçların çeviri eğitiminde kullanımını içeren çalışmaların yapılması alanın geleceği açısından önem taşımaktadır.

KAYNAKÇA

- All Languages Ltd. Translators & Interpreters. (2016). Machine translation basics. <https://www.alllanguages.com/documents/AllLanguagesMachineTranslation.pdf>
- Anderson, A. (2014). *Media, environment and the network society*. Londra: Palgrave Macmillan.
- Bowker, L. (2022). Machine translation literacy. <https://sites.google.com/view/machinetranslationliteracy/>
- Bowker, L. ve McBride, C. (2017). Précis-writing as a form of speed training for translation students. *The Interpreter and Translator Trainer*, 11(4), 259-279.
- Castells, M. (2004). Informationalism, networks, and the network society: A theoretical blueprint. Manuel Castells (Ed.), *The network society: A cross-cultural perspective* içinde (ss. 3-45). Cheltenham: Edward Elgar.
- Chesterman, A. (2016). *Memes of Translation: The spread of ideas in translation theory*, Amsterdam / Philadelphia: John Benjamins Publishing Company.
- Eisenstein, J. (2018). *Natural language processing*. MIT Press.
- Fantinuoli, C. (2023). EasyAI - Introducing artificial intelligence to the humanities. <https://easyai.uni-mainz.de/html/index.html>
- Hettige, B. ve Karunananda, A. S. (2011). Computational model of grammar for English to Sinhala machine translation. *International Conference on Advances in ICT for Emerging Regions (ICTer)* içinde IEEE, 26-31.
- Intento (2021). Independent multi-domain evaluation of machine translation engines. <https://try.inten.to/machine-translation-report-2021/>
- Johnson, M., Schuster, M., Le, Q. V., Krikun, M., Wu, Y., Chen, Z., Thorat, N., Viégas, F., Wattenberg, M., Corrado, G., Hughes, M., ve Dean, J. (2017). Google's multilingual neural machine translation system: Enabling zero-shot translation. *Transactions of the Association for Computational Linguistics*, 5, 339-351.
- Joseph, S. R., Hlomani, H., Letsholo, K., Kaniwa, F. ve Sedimo, K. (2016). Natural language processing: A review. *International Journal of Research in Engineering and Applied Sciences*, 6(3), 207-210.
- Khurana, D., Koli, A., Khatter, K. ve Singh, S. (2022). Natural language processing: State of the art, current trends and challenges. *Multimedia Tools and Applications*, 82, 3713-3744.
- Koehn P. (2005). Europarl: A parallel corpus for statistical machine translation. *Proceedings of the MT Summit* içinde, 5, 79-86.

- Lakew, S. M., Federico, M., Negri, M. ve Turchi, M. (2018). Multilingual neural machine translation for low-resource languages. *IJCoL. Italian Journal of Computational Linguistics*, 4(4-1), 11-25.
- Language Connections (2022). How interpreting services have changed at the olympics. <https://www.languageconnections.com/blog/blog-how-interpreting-services-have-changed-at-the-olympics/>
- Ma, Y. ve Tang, J. (2021). *Deep learning on graphs*. Birleşik Krallık: Cambridge University Press.
- Madill, W. (2022). Exploring natural language processing (NLP) in translation. <https://localizejs.com/articles/natural-language-processing-nlp/>
- Manning, C. ve Schütze, H. (1999). *Foundations of statistical natural language processing*. Massachusetts: MIT press.
- Millward, C. M. ve Hayes, M. (2012). *A biography of the English language* (3. Baskı). Amerika Birleşik Devletleri: Cengage learning.
- Ping, K. (2011). Machine translation. Mona Baker ve Gabriela Saldanha (Ed.), *Routledge encyclopedia of Translation Studies* (2. Baskı) içinde (ss. 162-169). Londra ve New York: Routledge.
- Randhawa, G., Ferreyra, M., Ahmed, R., Ezzat, O. ve Pottie, K. (2013). Using machine translation in clinical practice. *Canadian Family Physician*, 59(4), 382-383.
- Rothman, D. (2021). *Transformers for natural language processing: Build innovative deep neural network architectures for NLP with Python, PyTorch, TensorFlow, BERT, RoBERTa, and more*. Birmingham: Packt Publishing Ltd.
- Scott, M. L. (2015). *Programming language pragmatics*. New York: Morgan Kaufmann.
- Searle, J. R. (1980). Minds, brains and programs. *Behavioral and Brain Sciences*, 3, 417 – 457.
- Shreve, G. M. (2006). Integration of translation and summarization processes in summary translation. *Translation and Interpreting Studies. The Journal of the American Translation and Interpreting Studies Association*, 1(1), 87-109.
- Stenlund, S. (1990). *Language and philosophical problems* (1. baskı). Londra ve New York: Routledge.
- Vandermeulen, B. (2020). *Précis translation: Experiences, approaches and performance of third-year students of the bachelor in Applied Linguistics*. Universiteit Antwerpen: Lisans Tezi.

EXPANDED SUMMARY

Natural Language Processing (NLP) is a field of computer science and artificial intelligence that focuses on the interaction between computers and humans. NLP involves the development of algorithms and models that enable computers to understand, interpret, and generate human language. The goal is to provide machines with the ability to interpret and respond to human language as humans do, including tasks such as text classification, translation, sentiment analysis and many other applications. Today, the translation industry is heavily influenced by developments in NLP models, and translators use NLP-based systems to translate large volumes of text in a short time. The main NLP-based systems used by translators are machine translation software, post-editing tools, computer-assisted translation (CAT) tools, terminology management systems, automatic text summarization tools, translation management systems, translation quality assessment tools, computer-assisted interpreting (CAI) tools, speech recognition systems and chatbots. The goal of this study is to raise awareness of the NLP-based tools that are preferred for use in the translation process.

Machine translation is the most important area where NLP techniques are used. Today, most machine translation systems use neural networks trained on large datasets of parallel texts to learn how to translate between languages. This method produces more successful results than earlier approaches (rule-based, example-based and statistical machine translation systems), especially for resource-poor languages. Neural machine translation uses an approach called zero-shot translation to enable translation between arbitrary languages. Neural machine translation systems, which translate by taking into account relevant context and information content as well as linguistic features, are the best example of state-of-the-art NLP models. In addition, there have been remarkable developments in machine-assisted interpreting in recent years. Tools have been developed to assist the interpreter in both pre-interpreting and interpreting processes. However, the ability of machines to perform interpreting is still limited. Significant progress is being made in parallel with developments in artificial intelligence and natural language processing.

Another prominent field where NLP models are used is automatic text summarization. These tools help the translator understand the source text and extract key concepts. Translators can use automatic text summarization tools to create concise summaries of long and complex texts, which can then be translated more appropriately. Summaries can support translators in identifying key information, important concepts and the language register of the source text and produce more fluent and accurate target texts. Translators can also use these tools to gain domain-specific knowledge and terminology of a specific field. Today, there are automated tools that not only summarize the source text, but also translate it into the target language. AI-powered chatbots, new tools in the translator's workstation, can perform both extractive and abstractive summarization in the same language. They can also perform summary translation.

With the increasing demand for translation and the need to shorten translation processes, it has become almost a necessity for translators to use NLP-based tools. Therefore, developing an awareness of the concept of natural language processing and learning how these applications work will enable the translator to have a proper understanding of the limitations and capabilities of these systems. These tools can help translators produce higher quality translations, adapt to new technologies and improve their digital literacy. In today's global world, it is becoming increasingly important for translators to be familiar with the concept of NLP and its applications. This study attempts to analyze NLP-based applications used by translators, especially machine translation technologies, from a translation-oriented perspective. It is hoped that it will pave the way for further studies on the interdisciplinary relationship between NLP and translation studies.