İstanbul Üniversitesi Çeviribilim Dergisi
Istanbul University Journal of Translation Studies,
JTS 2025, (23): 73–93

https://doi.org/10.26650/iujts.2025.1711530
Başvuru | Submitted 01.06.2025
Revizyon Talebi | Revision Requested 13.08.2025
Son Revizyon | Last Revision Received 04.10.2025
Kabul | Accepted 11.10.2025

İstanbul Üniversitesi Çeviribilim Dergisi
Istanbul University Journal of Translation Studies

Araştırma Makalesi | Research Article

Açık Erişim | Open Access

Nöral Makine Çevirisinde Plastisite: Bilişsel ve Teknik bir Perspektif

Plasticity in Neural Machine Translation: A Cognitive and Technical Perspective

Derya Oğuz [1]

[1] Marmara Üniversitesi, İnsan ve Toplum Bilimleri Fakültesi, Mütercim ve Tercümanlık Bölümü, İstanbul, Türkiye

Öz

Nöroplastisite, insan beyninin değişken koşullara adapte olması ve uyum sağlaması konusundaki esnekliği ifade eden bir kavramdır. Yapay zekâ sistemlerinde ve nöral makine çevirisinde ise plastisite, çeviri çıktılarının doğruluğu ve bağlama duyarlılığı ile ilgili bir konudur. Bu çalışma, nöral makine çevirisinde plastisitenin nasıl sağlandığını ve bunu sağlayan mekanizmaların çeviri sürecindeki etkisini incelemeyi amaçlamaktadır. Plastisite olgusu, çalışmada veri temelli kavramsal analiz yaklaşımıyla ele alınmış, teknik boyutunun yanı sıra erek dilde anlamın yeniden yapılandırılması ve bağlamsal esnekliğin işlevselliği bakımından incelenmiştir. Bu doğrultuda dil modellerinde plastisiteyi sağlayan iki temel mekanizma ele alınmıştır: Birincisi hataya dayalı öğrenme süreci olan ve geri yayılım anlamına gelen backpropagation, ikincisi ise bir dil modelinin kullanıcı tarafından yeni girdilere uyumlanmasını içeren fine-tuning işlemi. Bu işleyişin arkasındaki mekanizma ise attention terimiyle ifade edilmektedir. Bu mekanizmalar öncelikle teknik özellikleriyle alanyazında araştırılmış, makine öğrenmesindeki plastisitenin insan beynindeki plastisite kavramından farkı ifade edilmiştir. Ardından bu mekanizmaların işlevselliği, örnek metinler, dikkat haritaları ve eğitilen bir dil modeli üzerinden yürütülen doküman analizi kapsamında betimleyici bir biçimde analiz edilmiştir. Elde edilen bulgular çeviribilimsel bakış açısından bilişsel süreçlerle ilişkilendirilmiş, model çıktılarının bağlamsal esneklik kapasitesi incelenmiştir. Bulgular, yapay öğrenme sistemlerinin nöral makine çevirisi yaparken yüzeysel anlam eşleştirmede başarılı olduğunu, bununla birlikte modelin işleyişinde istenmeyen yan etkilere de yol açabileceğini göstermiştir.

Anahtar Kelimeler dikkat · ince ayar · geri yayılım · nöral makine çevirisi · plastisite

Abstract

Neuroplasticity refers to the flexibility of the human brain in adapting to changing conditions. In artificial intelligence systems and neural machine translation (NMT), plasticity is related to the accuracy and contextual sensitivity of the translation outputs. This study aims to explore how plasticity is achieved in NMT and how its underlying mechanisms affect the translation process. The phenomenon is addressed through a data-driven conceptual analysis, focusing not only on its technical dimension but also on meaning reconstruction in the target language and the functional role of contextual flexibility. In this regard, two core mechanisms enabling plasticity are examined: backpropagation, an error-based learning process, and fine-tuning, which involves adapting a language model to new input by users. The underlying operation is driven by attention mechanisms. These processes are first investigated through the technical literature, highlighting how machine learning plasticity differs from human neuroplasticity. Then, their functionality is analyzed descriptively through a document analysis involving sample texts, attention maps, and an adapted language model. The findings are associated with cognitive processes from a translation studies perspective, showing that while NMT performs well in surface-level meaning matching, it may also produce unintended side effects within its operational logic.

Keywords attention · backpropagation · fine-tuning · neural machine translation · plasticity

Atıf | Citation: Oğuz, D. (2025). Nöral makine çevirisinde plastisite: Bilişsel ve teknik bir perspektif. *İstanbul Üniversitesi Çeviribilim Dergisi–Istanbul University Journal of Translation Studies*, (23), 73-93. https://doi.org/10.26650/iujts.2025.1711530

Sorumlu Yazar | Corresponding author: Derya Oğuz derya.oguz@marmara.edu.tr

Extended Summary

Neuroplasticity, originally used in neuroscience to describe the brain's ability to adapt and reorganize in response to changing circumstances, has gained increasing significance in the field of artificial intelligence (AI), particularly with the development of neural network architectures inspired by the human brain. In the biological context, plasticity refers to the brain's capacity to restructure itself, allowing healthy neurons to take over the functions of damaged ones. In AI, this adaptability is achieved through computational learning mechanisms and does not reflect structural reorganization in the biological sense. However, certain algorithmic processes in AI exhibit plastic-like behaviors, especially in natural language processing (NLP) and neural machine translation (NMT).

Recent advances in deep learning have enabled NLP models to adjust dynamically to varying inputs and linguistic contexts. In NMT, plasticity denotes the system's ability to detect and manage the structural and functional relationships between the source and target languages, influencing the contextual relevance and accuracy of the translations. This study aims to examine how plasticity is operationalized in NMT and how different computational mechanisms contribute to the adaptability and flexibility of translation outputs. It addresses the question of whether and how AI systems can simulate the flexible meaning-making processes of human translators. The research adopts a data-driven conceptual analysis method and uses descriptive document analysis across multiple levels of the translation process.

Two core mechanisms believed to enable plasticity in NMT are backpropagation and fine-tuning. Alongside these mechanisms, attention functions as an underlying system component, guiding the model's focus and contributing to contextual relevance. Each mechanism plays a distinct role in supporting the model's ability to adapt, learn, and restructure meaning representations. Backpropagation, or error-driven learning, is used to update a model's internal parameters by calculating and propagating the difference between the predicted output and the actual result. This mechanism is fundamental in training neural networks and is responsible for minimizing translation errors over time (Rumelhart et al., 1986).

In this study, the backpropagation mechanism was contextualized using linguistic data and real-word usage patterns from the DWDS corpus. Additionally, comparisons were made with the human brain's functioning based on Geoffrey Hinton's well-known argument that, while artificial neurons use symmetric weight updates, the human brain lacks this structure and operates non-hierarchically. This comparative lens sheds light on the fundamental differences in how biological and computational systems achieve adaptability.

Attention mechanisms guide the model to focus on the relevant parts of the input when generating translations. The self-attention mechanism, as proposed by Bahdanau et al. (2014) and further developed in transformer models by Vaswani et al. (2017), allows the system to assess the relative importance of every token in a sentence. This dynamic processing capacity plays a central role in enabling plastic responses to varying sentence structures and meanings. In the study, attention maps were used to visualize the alignment between the source and target language segments. A culturally and ideologically charged example, the German term "Gutmensch", was analyzed in German–English and German–Turkish translations. While the English output managed to convey a functional equivalence using the term "do-gooder," the Turkish output relied on a literal rendering (iyiliksever), which lacked socio-pragmatic adequacy. The attention maps showed that while the model allocated appropriate weights for English alignment, its sensitivity to contextual tone and connotation in Turkish was significantly weaker.

Fine-tuning involves adapting a pre-trained language model to a specific domain using task-relevant corpora. In this study, a bilingual Turkish–English model was fine-tuned using medical texts. The training corpus was adjusted by weighing medical texts ten times more than general texts to emphasize domain specificity. Attention weights and output quality were evaluated after every 1,000 steps, revealing that after 3,000 steps the model began to overfit, which reduced its contextual sensitivity. These findings highlight both the potential and limitations of domain-specific fine-tuning as a strategy for enhancing computational plasticity.

The study further integrates cognitive and translational theory by comparing machine outputs with expectations from translation studies. In particular, it draws on Baker's (1992) emphasis on the importance of context in meaning construction and Hatim & Mason's (2005) assertion that translation is not only a static transfer process but also a

dynamic activity shaped by contextual shifts. It also highlights the translator's interpretive flexibility, which involves cultural competence, pragmatic reasoning, and intuitive judgment — all of which contribute to human plasticity but are largely absent in current AI models.

In conclusion, this research shows that neural machine translation exhibits signs of adaptive plasticity through computational mechanisms, but still falls short of the human translator's contextual and interpretive abilities. While mechanisms such as backpropagation, attention, and fine-tuning enable AI models to simulate learning and alignment, true plasticity involves more than statistical patterning; it requires interpretive judgment rooted in a cognitive and cultural context. Therefore, future research and development efforts should prioritize hybrid systems that integrate human expertise and machine learning to achieve more accurate and context-sensitive translation outcomes.

Giriş

Plastisite sözcüğü, esasında beynin yeni durumlara adapte olması ve kendini geliştirmesi anlamında beynin esnekliğini ifade eden bir terim olarak sinirbilim/nörobilimde kullanılırken, yapay zekânın günlük hayatımızda daha çok yer almasıyla birlikte yapay zekâ alanında da önem kazanmaya başlamıştır. Yapay zekâ, bilindiği üzere yapay sinir ağlarıyla beyinden esinlenerek tasarlanmıştır ve bu nedenle beyin ile yapay zekâ sürekli karşılaştırılmaktadır. Söz konusu esinlenme, beynin işleyişi ile yapay sinir ağlarının işleyişi arasında birebir bir benzerlik olduğu anlamına gelmemekle birlikte, beynin plastisite özelliğinin yapay zekâ sistemlerinde ne ölçüde karşılık bulabileceği tartışmaya açıktır.

Beyinde plastisite, beyin dokusunun bazı etkenler sebebiyle yapısal ya da işlevsel bir değişikliğe uğraması durumunda (Anlar, 2013, s. 129) ya da genel anlamda herhangi bir sorun ortaya çıktığında sağlıklı olan nöronların diğer nöronların işlevini üstlenmesi ise karakterizedir. Beynin biyolojik plastisitesinden yalnızca bir sorun olduğunda söz edemeyiz. Beynin aynı zamanda çeşitli durumlara ve çevresel faktörlere bağlı yeni uyaranlara uyum göstermesi, beynin esnekliğiyle ilgili biyolojik bir niteliktir, diğer bir deyişle bir uyum gösterme yetisidir. Sözgelimi öğrenme eylemi beyinde çeşitli değişikliklere sebep olmaktadır (Açıkgöz ve Madi, 2013, s. 29). Öğrenme sırasında nöronlar arasındaki etkileşim ve elektrokimyasal işleyiş (Kılıç, 2024, s. 97) beyin dokusunu etkilemekte ve yeni bir yapılanma oluşturmaktadır. Bu yönüyle beyin, kendisini oluşturan dokunun bağlantısallığı ile dikkat çekmektedir. Bağlantısallık, beyin cerrahı Türker Kılıç'ın ifadesiyle beynin önemli bir özelliğidir. Beyin bağlantısallık temelinde çalışır (a.g.e., s. 80). Bu şekildeki işleyişin arkasında yatan mekanizma, bağlantısallık bütünsellik perspektifinden parça bütün ilişkisi içinde parçanın bütünü temsil edebilme yetisine dayanmaktadır. Bir bütünü oluşturan parçaların her biri bütüne ait özellikler taşır. Dolayısıyla beyinde herhangi bir bölümde temsil noktasında bir kısıtlılık ya da yoksunluk olduğunda diğer parçalar, eğer bütünü temsil etme bakımından güçlü ise söz konusu işlevi üstlenerek plastisiteyi sağlarlar. Bilgisayar sistemlerine ve yapay zekâya gelirsek, Kılıç'a göre eğer bütüne ait bir parça bütünü temsil etme noktasında yeterince güçlü ise bu mekanizma yapay zekânın plastisitesini sağlayabilir.[1] Bu bağlamda hatırlatmak gerekir ki beyin ve yapay zekâ sistemlerinde plastisite karşılaştırması, nörobilimsel bir iddia olmaktan ziyade kavramsal bir benzetme niteliği taşımaktadır ve elbette farklı bakış açılarıyla tartışılabilir. Kılıç'ın bakış açısıyla değerlendirildiğinde plastisite, yalnızca yapay zekâyı oluşturan teknik bileşenlerle değil; parçaların, sistemin bütüncül anlamda işleyişini yansıtabilecek ölçüde temsil işlevi görmesiyle ilgilidir. Diğer bir deyişle, yapay zekânın değişken koşullara uyum sağlayabilmesi, önceki deneyimlerinden elde edilen çıktıların ne ölçüde genellenebilir, esnetilebilir ve yeni bir bağlamda kullanılabilir olduğuna bağlıdır.

Doğal dil işlemede derin öğrenme algoritmalarıyla çalışan *Transformer* gibi güncel yapay zekâ mimarileri ile her gün yenilenen diğer yapay zekâ uygulamaları (*Gemini*, *BERT* vs.) bazı özellikleriyle belli ölçüde esneklik

[1] https://www.youtube.com/watch?v=jUZvlU1It5U&ab_channel=TekeTekBilim (01:12 dk.; 24.03.2025)

içermekte; öğrenme kapasiteleri, veri işleme biçimleri ve uyum gösterme davranışlarıyla insan beyni benzeri plastisite davranışları göstermektedir. Ne var ki henüz insan düzeyinde bir plastisite söz konusu değildir (en azından şimdilik). Yapay zekânın plastisitesi, halihazırda modelin öğrenmesi ve kendini güncellemesi ile sınırlıdır. Nöral makine çevirisinde ise plastisite, bir dil modelinin dil çiftleri arasındaki yapıyı ve ilişkiselliği anlaması ve elbette bu doğrultuda daha doğru çeviri çıktıları oluşması bakımından önemlidir.

Amaç ve Yöntem

Bu çalışma, nöral makine çevirisinde plastisitenin hangi mekanizmalarla sağlandığı ve bunun çeviri çıktılarında bağlamsal esneklik bakımından nasıl bir etki oluşturacağı temel sorusundan hareket etmektedir. Yapay zekânın plastisitesi, diğer bir deyişle esnekliği, özellikle nöral makine çevirisinde sistemin daha iyi öğrenmesi, dolayısıyla daha doğru sonuçlar verebilmesi demektir. Plastisite kavramı, çalışmada teknik boyutunun yanı sıra nöral makine çevirisinde erek dilde anlamın yeniden yapılandırılması ve bağlamsal esnekliğin işlevselliği bakımından ele alınmıştır. Bu doğrultuda nöral makine çevirisinde plastisiteyi sağlayan temel ve yardımcı mekanizmalar tespit edilmiş ve incelenmiştir.

Çalışma, plastisite olgusunu veri temelli kavramsal analiz yöntemiyle ele almış, örnek metin ve çeviri çıktıları üzerinden betimleyici ve yorumlayıcı bir yaklaşımla doküman analizi gerçekleştirilmiştir. Çalışmada yer verilen örnekler, nöral makine çevirisinde plastisiteyi sağlayan mekanizmalara karşılık gelecek biçimde ve bu mekanizmaların çeviri sürecindeki etkilerini gözlemlemeye mümkün kılabilecek biçimde amaçlı örnekleme tekniği ile seçilmiştir. Çalışmada öncelikle yapay zekâ sistemlerinde öğrenmeyi ve bağlamsal esnekliği sağlayan mekanizmalar alanyazın doğrultusunda açıklanmış; ardından bu mekanizmaların işlevselliği, metinler, dikkat haritaları ve tasarlanan dil modeli (*MultiMaCoCu*) çeviri çıktıları üzerinden analiz edilmiştir. Elde edilen bulgular, çeviribilimsel bakış açısından bilişsel kuramlarla ilişkilendirilmiş, modelin çıktılarının bağlamsal esneklik kapasitesi incelenmiştir.

Bilgisayar Sistemlerinde Plastisiteyi Sağlayan Mekanizmalar

Backpropagation- Geri yayılım

Yapay zekâda plastisite sağlamayı hedefleyen çeşitli mekanizmalar kullanılmaktadır. Bunlardan biri dil modelinin, hataları en aza indirmek ve daha doğru tahmin yaparak öğrenmek için kullandığı *backpropagation* algoritmasıdır. Bu algoritmayı, 1986 yılında arkadaşlarıyla yaptığı çalışmayla (Rumelhart ve ark., 1986) yapay zekânın gelişmesine sağladığı katkı nedeniyle bu yıl fizik alanında Nobel Ödülü alan Geoffrey Hinton geliştirmiştir. "Geri yayılım" anlamına gelen *backpropagation* mekanizması, nöral bir ağ sisteminde veriyi işledikten sonra katmanlar arasında geriye doğru yayılarak model eğitimi sürecinde tespit edilen hataları günceller. Aynı zamanda gradyan azalması (*gradyan descent*) yöntemi ile (Snelleman, 2016, s. 6) modelin doğruluğu artırılmaya çalışılır. Gradyan azalması yöntemi, modelin eğitiminde modelin öğrenmesini sağlayan ve geri yayılım sırasında ağırlıkları güncelleyerek *backpropagation* ile birlikte çalışan önemli bir mekanizmadır. Yapay sinir ağlarında katmanlar arasında sürekli matematiksel işlemler gerçekleşir ve buna göre her işleme/tahmine bir ağırlık değeri atanır. Nöral makine çevirisinde bir ağın eğitilmesi sürecinde bunun amacı kodlama ve kod çözme arasında (*decoding-encoding*) doğruluğun artırılmasıdır.

Her ne kadar yapay zekânın, insan beyninin çalışma biçimini referans aldığı söylense de mekanizmanın işleyişinde nöronal ağların yapısı dışında çok fazla benzerlik yoktur. *Backpropagation* mekanizmasını insan beyni ile karşılaştırmak için Hinton'un[2] "*Can the brain do backpropagation*?" (İnsan beyni geri yayılım yapabilir mi?) başlıklı seminerden faydalanabiliriz. Hinton, bu seminerde insan beyninin *backpropagation*

[2] https://www.youtube.com/watch?v=VIRCybGgHts&ab_channel=StanfordOnline (24.03.2025)

mekanizmasını tam olarak uygulamadığını belirterek çeşitli gerekçeler sunar. Bunlardan ilki, insan beyninin doğrudan hata sinyali işleyen, hesaplayan ve geriye doğru yayarak düzenleme yapan mekanik bir ağ yapısına sahip olmamasıdır. İnsan beyni bu mekanizmadan farklı hiyerarşik olmayan bir işleyişle çalışır. Yapay nöronlar ise, tahmini çıktı ile olması gereken çıktı arasındaki farkı matematiksel olarak hesaplayarak ağın ağırlıklarını günceller. Hesaplama, gradyan azalması yöntemine dayanır ve mekanik bir işlemdir. Oysa insan beyninde olması gereken ile tahmin arasında sürekli bir hesaplama söz konusu değildir. Biyolojik/kortikal nöronlar paralel işlem yapabilen bir niteliğe haiz iken yapay nöronlar hiyerarşik ve mekanik çalışırlar.

Hinton, insan beyni ile yapay nöronları karşılaştırdığı seminerde beynin mekanik bir biçimde geriye yayılım yapmamasını birkaç farklı gerekçeyle detaylandırır. Yapay nöronlar bir veriyi işlerken katmanlar arasında ilerlerken (*feed forward*) matematiksel değerler ile hesaplamalar yaparak sürekli nitelikte aktivasyon değerleri oluşturur. Oluşturulan değerler, farklı değerlerdedir, süreklilik arz eder ve bir sonraki katmana iletilir. Kortikal nöronlarda ise nöronlar arasında bu şekilde bir aktivasyon değeri hesaplanmaz. Nöronlar arasında bir aktivasyona/ateşlemeye 1 değeri verirsek ateşlememe olmamasına 0 diyebiliriz. İnsan beyninde nöronlar arası ateşleme (*spike*) belli bir dereceye kadar uyarılmış olmak şartıyla ya gerçekleşir ya gerçekleşmez (ya hep ya hiç ilkesi, Cüceloğlu, 2015, s. 60); bu yönüyle kortikal nöronlar yapay nöronlardan farklıdır. Yapay nöronlarda sürekli nitelikte bir sinyal/değişken (*real valued*) söz konusuyken, insan beyni böyle çalışmaz.

Hinton'un öne sürdüğü diğer bir gerekçe, yapay nöronların katmanlar arasında hem ileri yayılarak ilerlerken bir çıktı oluşturması hem de geriye yayılarak bir hata tespiti sonucunda çıktı oluşturması karşısında insan beyninin çift yönlü sinyal işlememesidir. Sonuncusu ise yapay nöronların aksine insan beyninde kortikal nöronların simetrik bir geribildirim işleyişinin olmamasıdır. Bu doğrultuda diyebiliriz ki insan beyninin esnekliği ve plastisitesi yapay zekâdan oldukça farklıdır. Temelde beynin işlevi taklit edilmeye çalışılsa da mekanizmaların farklılığı insan beyni düzeyinde bir plastisiteye erişimin henüz mümkün olmadığını göstermektedir. Eksponansiyel (üstel) bir hızla ilerleyen teknoloji sayesinde yapay zekâda hedeflenen ölçüde plastisite, önümüzdeki günlerde mümkün olabilecek olsa da mevcut durum böyle betimlenebilir. İnsan beyni sinaptik bağlantılarını, etrafında cereyan eden olumlu ya da olumsuz sonuçlara göre biçimlendirerek işler. Buna göre geribildirimler yoluyla sinaptik bağlantılar güçlenebilir ya da zayıflayabilir. Geri bildirimi, kabaca *backpropagation* mekanizmasına benzetebiliriz ancak yukarıda sayılan farklılıklardan dolayı beyindeki ve yapay zekâdaki işleyiş tam olarak bağdaştırılamaz. İnsan beyni kendi karmaşıklığı içerisinde şüphesiz plastisite bakımından farklıdır. Beyinle ilgili yapılan çalışmalar[3] incelendiğinde anlaşılan o ki beyin, henüz işleyiş bakımından gizemini korumaktadır. Bilgisayar teknolojilerinin mevcut kapasitesi insan beyninin tam olarak simüle edilmesi için hala yeterli değildir.

Dikkat çekici olan, yapay zekânın *backpropagation* gibi bir mekanizmanın yardımıyla insan benzeri çıktılar üretebilmesine karşın insan beyninin tüm karmaşıklığıyla daha esnek ve sezgisel sonuçlara ulaşabilmesidir. İnsan doğası gereği bunu zaten yapar ve bu, çok da şaşırılacak bir şey değildir. Her gün otomatik olarak yaptığımız ve sıradan gördüğümüz bu süreçlerin arkasında yatan mekanizmayı farkında olmadan işletiriz. Beynin nasıl çalıştığını anlamak, şüphesiz yapay zekânın işleyişine ve gelişimine dair bilgi verebilir. Bu bağlamda konuyla ilgili çalışmaları takip etmek, anlam üretimiyle doğrudan ilişkili çeviri ediminde gerçekleşen bilişsel süreçlerin anlaşılmasına ve insan benzeri esneklik ve yaratıcılığı taklit eden sistemlerin geliştirilmesine hizmet edebilir.

[3]Avrupa'da gerçekleştirilen "İnsan ve beyin projesi" (bkz. https://www.humanbrainproject.eu/en/) ve Amerika'da gerçekleşen "İnsan konnektom" projesi (bkz. https://sinirbilim.org/konnektom/) bize bununla ilgili bilgi verir. Yaklaşık 100 milyar nörona sahip beynin sırları beyin tam olarak bütünüyle simüle edilemediği için henüz çözülememiştir ancak bu projelerle özellikle teknolojik gelişmelere ve yapay zekâya zemin hazırlayabilecek oldukça önemli veri sağlanmıştır. Beyin alanında yapılan en güncel çalışma (Schlegel ve ark., 2024) bir meyve sineğine ait beynin haritalamasıdır. Bu çalışmaya göre yaklaşık 140 bin nörona ve 50 milyona yakın sinaptik bağlanıtısına sahip sineğin beyni, elektron mikroskobu ve yapay zekâ yardımıyla üç boyutlu olarak modellenmiştir.

Backpropagation mekanizması da esnekliği sözü edilen sistemlerden biridir ve henüz gelişmektedir. Zira güncel bir çalışma (Dohare ve ark., 2024), derin öğrenme ve geri yayılım algoritmalarının yapay zekânın temelini oluşturduğunu belirterek, bu algoritmaların sürekli öğrenme ve esneklik bakımından henüz teknik olarak yeterli olmadığını ileri sürmüştür. Araştırma ekibi, çeşitli veri setleri üzerinde çalışarak yapay zekânın plastisite kaybını (*loss of plasticity*) göstermeye çalışmışlardır. Çalışmada plasitisite kaybıyla ilgili sözü edilen diğer bir kavram *backpropagation* mekanizmasıyla birlikte çalışan "gradyan inişi" dir (*gradient descent*). Gradyan inişi, sistemin ne kadar esnek olabileceğini belirleyen bir bileşendir. Gerçekleşen geriye yayılım, hataların yapay sinir ağı katmanlarında nasıl yayılacağını hesaplarken, gradyan inişi bu ağırlıkları temel alarak ağırlıkları günceller. Çalışmada ortaya konan önemli bir nokta, esnekliğin artması için gradyan inişine bağlı olmayan rastgele, düzenli ve öngörülebilir olmayan bir mekanizmanın gerekliliğidir.

Nöral makine çevirisinde genel anlamda bakıldığında kullanılan mimarilerin çoğu gradyan inişi temellidir (*Transformer; LSTM- Long Short Term Memory*).[4] Gradyan inişine alternatif olarak araştırılan ve kullanılan farklı yöntemler de mevcuttur. Sözgelimi gradyan dışı ve deneme-yanılma yoluyla geri bildirim veren bir mekanizmaya sahip, pekiştirme yoluyla öğrenme anlamına gelen *RL-Reinforcement Learning* (Alleman, Atrio ve Popescu-Belis, 2024) mekanizması, plastisite bakımından sorgulanan bir konudur.[5] Araştırmaya göre, uzun dönemde sürekli öğrenmeye maruz kalan mekanizmanın esnekliğini kaybettiği görülmektedir. Bir süre sonra mekanizma, yeni bilgilerle karşılaştığında yenilere uyum sağlamaya çalışırken önceki bilgileri birden silebilmekte, böylelikle sistemde yıkıcı bir etkileşim gerçekleşmektedir (*catastrophic-interference*).

Makine çevirisinde plastisitenin sürdürülmesi konusunda yapılan çalışmalar her geçen gün ivme kazanmaktadır. Her yeni çalışma, önceki çalışmaları ve yapay zekâ sistemlerini anlamak bakımından birbirini tamamlayıcı niteliktedir. Çeşitli mimarilerin bir metni nasıl işlediği ve plastisiteyi nasıl sağlamaya çalıştığı üzerine bulgular, bu sistemlerin çeviri süreçlerinde nasıl işlediğine ve nasıl birbirinden farklılık gösterdiğine dair bilgi vermektedir. Nöral makine çevirisi bağlamında plastisite, çeviri çıktıları üzerinden gözlemlenebilir biçimde karşımıza çıkmaktadır. Bu çerçevede nöral makine çevirisinde plastisitenin nasıl tezahür ettiğini somut örnekler üzerinden görmek yerinde olacaktır.

Nöral Makine Çevirisinde Plastisite

Backpropagation mekanizması, yukarıda söz ettiğimiz üzere yapay sinir ağlarında öğrenme sürecini yöneten bir algoritma olarak nöral makine çevirisinde karşımıza çıkmaktadır. Bir çeviri örneğinde makinenin yaptığı çevirinin hatalı olması neticesinde, hata bütün katmanlar arasında geriye doğru yayılarak ağırlıkları günceller. Ağırlık güncellemesi sözcükler arasındaki bağlantıların/bağlantısallığın güncellenmesini sağlar ve model buna göre öğrenerek bir sonraki çeviride daha doğru karar verir. Sözgelimi Almancada günlük dilde "Naber?" anlamında kullanılan "Was geht?" ifadesi çeviride "Ne oluyor/ Ne gitti/Ne gidiyor?" gibi sözcük düzeyinde karşılıklar bulabilir. Çeviri modeli, bağlamı tanıyıp dilin kültürel örüntülerini öğrendiğinde elbette sonuç farklı olur ve uygun karşılıklar üretilebilir. Nöral makine çevirisinde bu ifade için bağlama uygun karşılıkları alabilmek, sistemin dilsel veriyi işlemesinin yanı sıra bağlamı tanıyan bir plastisite kapasitesine sahip olmasına bağlıdır. Ne zaman ki modele bu ifadeyle ilgili doğru karşılık verilir ve bu birçok kez tekrarlanırsa, sistem yaptığı çeviriyi bir hata (*loss*) olarak kabul edip doğru çeviriyi esas alır. Hata, modelin her bir katmanına yayılarak sözcükler arası ilişkilerde ağırlıklarını güncelleyerek daha doğru çeviri yapmayı öğrenir. Bu süreç, modelin farklı bağlamlarda zamanla nasıl tepki vereceğini öğrenmesini sağlar. Böylece sistem katı bir ezber üzerinden değil, esnek bir öğrenme yapısı üzerinden işlemeye devam edebilir. Plastisite

[4] Ayrıntılı bilgi için bkz. Zhao, Y., Zhang, J., ve Zong, C. (2023); Chen, N. (2024).

[5] İlgili çalışma için bkz. Abbas, Z., Zhao, R., Modayil, J., White, A. & Machado, M. C. (2023, November).

burada sistemin her yeni veriyle birlikte kendini güncelleyebilmesi ve yeni durumlara daha doğru ve uygun karşılıklar üretebilme becerisini ifade eder.

Peki insan beyninde plastisite nasıl sağlanır? Normalde Almanca günlük dile çok hâkim olmayan biri benzer bir çeviri hatasına düşse insan beyni bu hatadan nasıl öğrenir? İnsan beyninde daha önce de söz ettiğimiz üzere *backpropagation* gibi bir mekanizmanın tam olarak karşılığı yoktur. Yapılan çalışmalar insan beyninin yeni sinaptik bağlantılarla beyin bilgilerini güncellediğini göstermektedir. Bu noktada denilebilir ki insan beyni bir hatayı fark ettiğinde ve doğrusunu öğrendiğinde sinaptik bağlantılarla beyin bağlantısallığı diğer bir deyişle konnektomu[6] (Kılıç, 2024, s. 55) güncellenmiş olur. Nörobilim araştırmacılarının açıkladığı deneyime bağlı öğrenme (*experienced-dependent learning*) yaklaşımına göre yaşanan deneyimin türü, beyni etkilemektedir ve beynin plastisitesi buna göre değişmektedir (Wenger ve Löwdén, 2016, s. 171).

Beynin nasıl öğrendiği ve hatayı nasıl işlediği bir tarafa, bu noktada yorum yapabilmek için dilin nasıl işlendiğini bilmek gerekir. Dilin beyinde nasıl işlendiğine dair daha önceki bilgimiz, beyinde konuşma ve anlamaya ilişkin belli merkezlerin (*Broca Alanı* ve *Wernicke Alanı*) olduğu yönünde idi (Kerimoğlu, 2022, s. 34). Ancak güncel çalışmalar bildiklerimizin sınırlı olduğunu göstermekte, beyinde dil ile ilgili aktif alanların yanı sıra farklı bir işleyişin olabileceğine işaret etmektedir. Zira Minnesota Üniversitesi'nde aralarında görüntüleme alanında dünyanın en saygın bilim insanlarından biri olan Prof. Dr. Kâmil Uğurbil'in bulunduğu beyin fMRI gerçekleştiren araştırmacılar (Hinke ve ark., 1993), sözcüklerin beyinde bağlantısallıkları ile birlikte farklı yerlerde konumlandığını göstermektedir.

Bu doğrultuda tekrar yukarıda verdiğimiz "was geht?" çeviri örneğine dönersek, beyinde "was" sözcüğü ile "gehen" sözcüğü ayrı, bunun yanı sıra "Was geht" kalıbı yeni bir sinaptik bağlantı ile farklı bir renkle ifade edilebilir. "Was geht?" örneği üzerinden diyebiliriz ki; nöral makine çevirisinde bağlama dayalı yorumlama yetisi, dilsel çözümlemenin yanı sıra sistemin öğrenme mimarisi ve plastisitesi ile ilgilidir. Bu tür kültüre özgü ifadelerin nöral makine çevirisi ile insan beyni tarafından işlenişi arasında farklılıklar söz konusudur. Beyin deneyime dayalı ve sezgisel bir biçimde hareket ederken, yapay zekânın örüntü üzerinden istatistiksel bir tercih yaptığı görülmektedir. İki düzenek arasında bir benzerlik kuracak olursak yapay zekâ ile beynin işleyişinde benzer noktanın, kabaca iki düzeneğin hatalardan öğrenmesi olduğunu söyleyebiliriz.

Plastisiteyi yine çeviriye dair somut bir örnek üzerinden açımlayalım. Almancada polisemik bir yapısı olan ve anlam ayrıştırma bakımından plastisiteyi ölçmeye zemin hazırlama niteliğinde temel anlamı "yöntem" olan "Verfahren" sözcüğü bu bağlamda uygun bir seçimdir. Sözcüğün Almanca dijital sözlükteki (*DWDS- Digitales Wörterbuch der deutschen Sprache*)[7] kullanımları ve çevirileri bağlamsal esneklik bakımından örnek teşkil edebilecek niteliktedir. DWDS plaftormunda sözcüğün kullanım biçimlerine dair dizi ya da filmlerden altyazı örnekleri verilmiştir. Bunlardan birisi "Wollt ihr mir jetzt dabei helfen, herauszufinden, warum dieses Verfahren nicht funktioniert?" cümlesidir. Bu cümlede sözcüğün anlamı açıktır ve "Bu yöntemin/sürecin neden işe yaramadığını bulmama yardım etmek ister misin?" biçiminde çevrilebilir. Ancak bu sözcüğün hukuki bağlamında kullanımıyla ilgili "Das Verfahren wurde eingestellt." gibi bir örnekte "bir davanın düşürülmesi, yargılama sürecinin sona ermesi" gibi bir anlamdan söz edilmektedir. Nöral makine çevirisinde sistem, daha önce işlenmiş verilerden öğrenerek çok katmanlı anlam yapılarını ayırt etmeyi öğrenmektedir. Dil modelinin plastisite kapasitesi yeterli ise çeviri çıktısı da bu doğrultuda olacaktır.

Sözgelimi "The Following" dizisinin 2014 yılında yayınlanan "Sacrifice" (Sezon 2, Bölüm 12) bölümünde geçen "Alle Mitglieder der Gemeinschaft haben mehrfach dieses Verfahren durchlaufen" (*undergo a procedure*) altyazısı, dizinin ana karakterlerinden Joe Carroll ve takipçilerinin organize bir tarikata katılma

[6] Öğrenme, beyinde nöronlar arasında yeni bağlantısallıklar kurma sürecidir. Öğrenme sırasında beynin hangi alanı daha çok kullanılıyorsa o bölgedeki enformasyon yollarında (nörozihin=konnektom) bir nicelik ve karmaşıklaşma saptanmaktadır.

[7] https://www.dwds.de/wb/Verfahren

sürecini anlatan bir sahnede yer alır. Burada "Verfahren" sözcüğü "bir tür sınav ya da ritüel, prosedür, süreç, uygulama" anlamında kullanılmıştır ve şöyle çevrilebilir: "Topluluğun tüm üyeleri bu prosedürden birkaç kez geçmiştir." Makine çevirisinde doğru bağlamda çeviri çıktısı elde edebilmek için dil modelinin bu bağlamı öğrenmiş olması gerekir. Dil modeli yeterince öğrenmemiş ve plastisitesi yeterli değilse bu cümlede "Verfahren" sözcüğü "yöntem" olarak çevrilebilir.

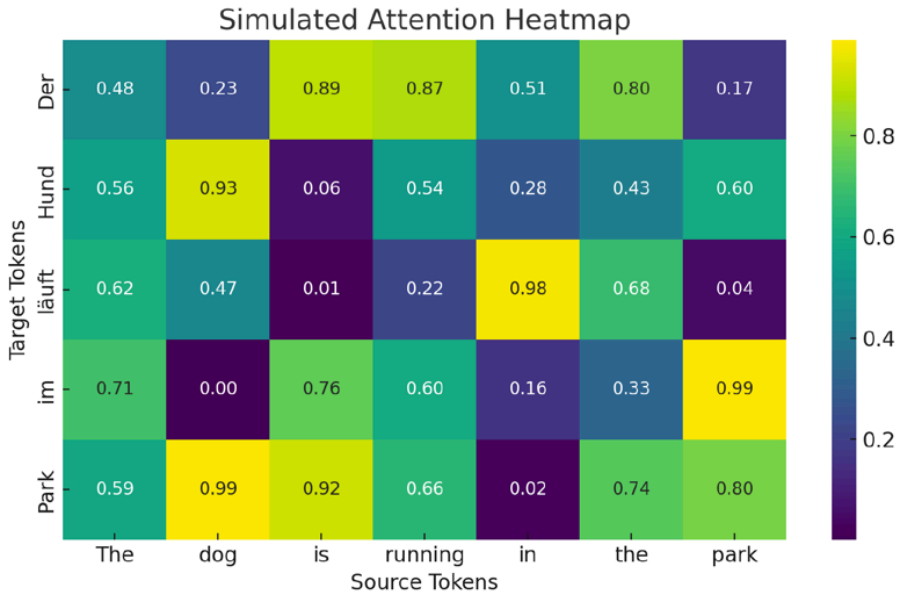
Yukarıda sözü edilen bağlamsal farklılıkların nöral makine çevirisinde doğru biçimde yönetilebilmesi ve sistemin istenen ölçüde bir plastisiteye sahip olması, kullanılan mimaride sistemin öğrenmesiyle ve belleğiyle ilişkilendirilebilir. Dil modelinin doğru anlamı yakalayabilmesi, sağlanan geri bildirimler yoluyla sistemin ağırlıklarının güncellenmesine bağlıdır. Nöral ağlarda katmanlar arasında işleyerek geri bildirim sağlayan *backpropagation* mekanizması, yukarıda söz ettiğimiz üzere hata güncellemeyi ve öğrenmeyi mümkün kılar. Öğrenmenin verimliliği ise plastisiteyi olumlu anlamda etkiler. Bu süreçte nöral ağların genel işleyişinin temel bileşenlerinden biri olan ve sistemin hangi girdiye ne düzeyde odaklanacağına karar veren "attention" adında başka bir mekanizma daha girer devreye. Dikkat anlamındaki mekanizma, yapısal olarak tüm sistemin çalışmasına entegre bir unsur olsa da karmaşık yapılarda sistemin doğru çıktı için neye odaklanması gerektiğine karar vererek dolaylı anlamda plastisiteye katkıda bulunur.

Attention-Dikkat

Yapay zekânın plastisitesiyle ilişkilendirebileceğimiz diğer bir nitelik nöral ağ sistemine entegre olarak çalışan *attention* (dikkat) mekanizmasıdır ve bu mekanizma geliştirilerek *self-attention* olarak adlandırılmıştır. Büyük dil modellerinde en güncel mimarilerin kullandığı *self-attention* (Bahdanau ve ark., 2015, s. 4; Vaswani ve ark., 2017, s. 2-10), her bir girdi metninde tüm sözcüklerin/ögelerin birbiriyle ilişkisini inceler, temel ögeleri belirler ve bağlamı değerlendirir. Bunu yaparken bilgiyi dinamik olarak işler, her yeni girdide bağlantısallıkları göz önünde bulundurur; birbirine uzak sözcükler arasındaki ilişkiyi değerlendirir ve uzun vadede daha doğru çeviriler yapmayı amaçlar. Aynı zamanda sözcükler arasında önemli olan sözcükleri bağlama göre belirler ve bu bilgileri sürekli güncelleyerek adapte olur ve öğrenmeye devam eder.

Dil modelleri bir çeviride modelin bağlamda neye göre karar verdiğini ve en çok neye dikkat gösterdiğini açıklanabilir yapay zekâ kapsamında "*attention map*" adı verilen dikkat haritaları ile görüntüleyebilir (Vig, 2019). Dikkat haritaları ile yapılan çalışmalar nöral makine çevirisinde dil modelinin davranışını analiz etmek için önemli araçlardır. Dikkat haritaları, teknik bir çıktı olmanın yanı sıra aynı zamanda dil modelinin bağlamı ne ölçüde içselleştirdiğini gösteren bir gösterge olarak değerlendirilebilir. Dikkat haritaları aynı zamanda açıklanabilir yapay zekâ (*XAI-explainable artificial intelligence*) alanı için de önemli veriler sunar. Aşağıda kaynak metin İngilizceden erek dil Almancaya nöral makine çevirisi ile gerçekleştirilmiş bir çıktıya ilişkin dikkat haritası örneği görülmektedir (Görsel 1). Dikkat haritası, kaynak ve erek metin arasında bağlantı kurarak metinler arası sözcük, sözdizimi ve anlam ilişkilerini inceleyerek bir tür eşleştirme gerçekleştirir. Gerçekleştirilen eşleşmede sözcükler arasındaki ilişkinin niteliğine göre sayısal bir değer atayarak ağırlıklar oluşturur. Ağırlıklar, bunları gösteren renklere göre açıktan koyu renge bir skala oluşturduğu için ısı haritası olarak da anılabilir.

Görsel 1
Dikkat Haritası



İnsan beyninde genel anlamda dikkat mekanizması, başta alın kısmında yer alan beynin prefrontal korteks bölgesi, parietal alan ve bazı diğer bölgelerin etkileşimi ile işleyen karmaşık bir süreçtir (Corbetta ve Shulman, 2002; Posner ve Petersen, 1990, s. 25/34). Beyin bu sistem sayesinde neye odaklanacağına karar verebilir. İnsanda dikkat mekanizması çeviri sürecinde de devreye girer. Yapılan bir çalışmada (Price ve ark., 1999, s. 2221) çeviri sırasında beynin aktive olan bölgeleri görüntülenmiş, dilin işlenişi sırasındaki dikkat ve karar verme ile ilgili ön bölgede (*anterior singulat korteks*), duygu ve otomatik tepkileri yöneten subkortikal yapılarda (serebral korteksin altındaki bölgeler) ve *Broca* alanında aktivasyon gözlemlenmiştir. Bulgular, çeviri sürecinde bilişsel düzeyde yoğun bir dikkat ve seçme işlemi gerçekleştiğini göstermektedir. Bu yönüyle insan beyni ile yapay zekâ sistemlerindeki dikkat mekanizmasında işlevsel bir benzerlik kurabiliriz. Beyindeki mekanizmanın işleyişine dair çalışmalar devam etmektedir ancak mekanizmanın daha karmaşık bir yapıda işlediği ve doğal bir plastisite sergilediği söylenebilir.

Nörobilim çalışmalarıyla paralellik gösterecek biçimde çeviribilim alanında çeviri sırasında dikkat, bellek ve karar verme gibi bilişsel süreçlerin incelendiği çalışmalar mevcuttur. Finlandiyalı çeviribilimci Riitta Jääskeläinen (2011), çevirmenin çeviri sırasındaki bilişsel akışını "sesli düşünme protokolleri" (TAPs) uygulamasıyla gözlemlemeyi önermiştir. Sesli düşünme sırasında çevirmenin düşüncelerini içeren sözlü ifadeleri kaydedilmekte ve böylelikle çeviri edimine yönelik ampirik veri toplanmaktadır. Her ne kadar kaydedilen veriler, beynin iç işleyişine yönelik bilgi vermese de bilişsel psikolojiden esinlenerek yürütülen uygulamalarda çevirmenin zihinsel eforuna, dikkat ve karar verme süreçlerine dair inceleme ve gözlemler gerçekleştirilmiştir. Barselona Üniversitesi bünyesinde kurulan çeviri edincine yönelik çalışmalarıyla bilinen PACTE araştırma grubu (*Process of Acquisition of Translation Competence and Evaluation*), yaptığı bir araştırmada çevirmenin karar verme sürecini incelemiş, çevirinin bilişsel bileşenlerini araştırmıştır (bkz. PACTE, 2017). Sözü edilen araştırmalar insan çevirmenin bilişsel kapasitesi ve çeviri sürecinde beynin plastisitesine yönelik araştırmalar kapsamında değerlendirilebilir.

Attention ağırlıklarıyla "Gutmensch" analizi

Almancada 2015 yılında bir jüri tarafından yılın en uygunsuz sözcüğü seçilen "Gutmensch" ifadesinin geçtiği birkaç bağlamı inceleyelim. Metnin seçilmesinin sebebi, metinlerde geçen *Gutmensch* sözcüğünde

anlamın, sözcük düzeyinde değil, sözcüğün ötesinde aranması gerektiği; bunun yanı sıra metnin söylemsel bir nitelik barındırması, metinlerarası göndermeler, imalar içermesi ve çeviride pragmatik düzeyde yeniden inşa edilme zorunluluğudur. Hatim ve Mason (2005, s. 5) çeviriyi, metin üreticilerinin dünyaya dair bakış açılarını alımlayıcılara ileten söylemsel göstergelerin aktarımı olarak görür ve anlamın sözcüklerin ötesinde kurulduğunu vurgularlar. Seçilen metin, bu doğrultuda anlamın pragmatik düzeyde kurulduğu bir yapıya sahiptir ve çeviri süreci de yorumlayıcı bir esneklik gerektirir. Baker'ın (1992, s. 217-219) belirttiği üzere çeviride pragmatik düzeyde bir eşdeğerlik yakalanması önemlidir. Bir metnin anlam bütünlüğü, ancak metinde sunulan bilgi ile okuyucunun kendi bilgisi ve dünya deneyimi arasındaki etkileşimin sonucunda sağlanabilir ve bu etkileşim yaş, cinsiyet, ırk, milliyet, eğitim, meslek, siyasi ve dini eğilimler gibi çeşitli faktörlerden etkilenir. Baker (1992, s. 238), anlamın belirlenmesinde ayrıca bağlamın önemini vurgulayarak bağlamın, ondan makul bir biçimde çıkarılacak birçok ima içerebileceği olasılığına dikkat çeker. Bu noktada insan çevirmenin metne yaklaşımı ve anlamı teşhis etmeye çalışması ile nöral makine çevirisinin metni çözümlemesi arasındaki olası fark devreye girer. İnsan çevirmen böyle bir çeviri görevinde edinçlerini kullanarak bağlamsal, kültürel ve sezgisel yoruma dayalı bir süreç işletirken, nöral makine çevirisinde büyük ölçüde veriye ve örüntüye dayalı bir dizi işlem ve eşleştirme gerçekleşir. Aşağıda "Gutmensch" metni üzerinden yer alan analiz, bir taraftan iki yaklaşım arasındaki farkı görünür kılmayı amaçlarken diğer taraftan sistemin plastisitesini sağlayan mekanizmaların bağlamı nasıl işlediğini ortaya koymayı amaçlamaktadır.

"Gutmensch" metni ideolojik referansları nedeniyle belirtildiği üzere özellikle seçilmiştir. Metinde geçen "Gutmensch" sözcüğünün ilginç olan tarafı, yardımseverliğin, sosyal adaletin karalanmasına, siyasette nefret söyleminin ve kutuplaşmanın artırılmasına, siyasi tartışmaların rasyonellikten uzaklaşarak kişisel bir saldırı haline gelmesine zemin hazırlamasıdır. "Gutmensch" sözcüğü Almancada siyasi bağlamlarda ve popülist söylemlerde siyasilerin rakiplerine karşı kullandıkları bir ifadedir. İronik, suçlayıcı, aşağılayıcı ve eleştirel bir bağlamda negatif anlamda kullanıldığı göze çarpar. Bunun için konuda biraz daha derinleşip çeşitli makaleleri taradığımızda ve sosyal medya analizi yaptığımızda konunun ayırdına varırız.[8] Genellikle samimiyetsiz, çıkarcı davranışları olan, kendi görüşü dışındakilere hoşgörüsüz, belli bir ideolojiye körü körüne bağlı kişileri tanımlamak için kullanıldığı görülür. Sözgelimi bağlam vermediğimiz "Gutmenschen sind keine guten Menschen." cümlesi GPT-4 modeli tarafından Türkçeye "İyiliksever insanlar iyi insanlar değildir." biçiminde çevrilmiştir. Model, cümleyi İngilizceye ise "Do-gooders are not good people." olarak çevirmiştir. İngilizce çeviride dikkati çeken, "Gutmenschen" sözcüğünün, görece Almancadaki anlamsal ve işlevsel karşılığına uygun biçimde verilmesidir. Türkçe çeviride "iyiliksever" sözcüğü, Almancadaki işlevine oranla sözcük düzeyinde bir çeviri olarak kabul edilebilir ancak işlevsel değildir. Zira "Gutmensch" sözcüğü Almancada özellikle siyasi bağlamlarda rakibini küçümsemek ve itibarsızlaştırmak için kullanılan kavramsal bir silah işlevi görmektedir. Bu bağlamı prompt olarak sunduğumuzda, model yaptığı çeviriden bağımsız olarak anlama ilişkin şu açıklamayı vermiştir:

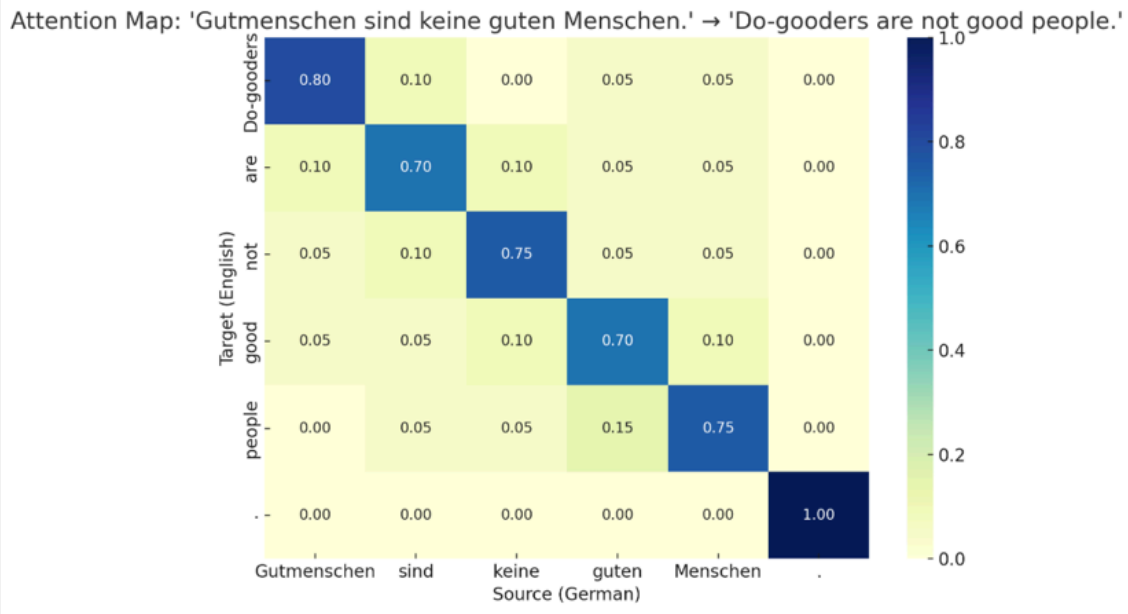
> Bu bağlamda Gutmenschen ifadesi, ironik ya da aşağılayıcı bir anlamda kullanılarak, "iyi görünmeye çalışan" kişileri ifade eder. Cümlede, bu kişilerin aslında "iyi insanlar" olmadığı belirtiliyor.

Açıklamanın ardından yapılan çeviri ile ilgili bir *attention map* oluşturması istenmiş (bkz. Görsel 2), metni bu yöntemle nasıl çözümlediği sorulmuştur.

[8] https://www.ndr.de/kirche/Gutmensch-Ein-Wort-dass-sich-radikal-veraendert-hat,unwort140.html
https://www.deutschlandfunkkultur.de/gutmenschen-eben-mal-die-welt-retten-100.html
https://www.zeit.de/gesellschaft/2015-04/altruismus-gutmensch-motive
https://www.dw.com/de/woke-gutmenschen/a-68089581

Görsel 2

İngilizce çeviri için dikkat haritası



Self-attention mekanizmasının nasıl işlediğini sorduğumuzda ise GPT-4 modelinin ChatGPT arayüzü üzerinden verdiği cevaplar ilginçtir. "Gutmenschen sind keine guten Menschen." cümlesinde "Gutmenschen" sözcüğünün özne olduğunu, "sein/sind" fiilinin ise yüklem olduğunu öncelikle belirlemiştir. "Gutmenschen" ile "guten Menschen" arasındaki ilişkiyi "sein" fiili kurmaktadır. *Self-attention* mekanizması ile "Gutmenschen" öznesinin "guten menschen" ifadesi ile karşılaştırıldığını ve niteliksel bir karşılaştırma yapıldığını fark etmiştir. "Sein" fiili her iki sözcüğe *attention* mekanizması ile, diğer bir deyişle dikkatle bağlanır. "Keine" sözcüğü, "Gutmenschen" ile "guten Menschen" arasındaki olumsuzluk bağını kurmuştur. *Self-attention* mekanizması, "keine" sözcüğünün çekimli fiil olan "sind" ve nesne olan "guten Menschen" ile nasıl ilişkili olduğunu öğrenir. Buraya kadarki açıklamada son cümlede nesne olarak belirtilen "guten Menschen" ifadesi ile ilgili bilgi, dil bilgisel yapıların açıklanması ve cümle bilgisi açısından tam olarak doğru değildir. Bu ifade, aslında "sein" fiili ile birlikte eyleme dahil edilebilecek "Prädikativ" (ek eylem) yapısıdır.

Bu noktada dikkati çeken, "Prädikativ" cevabının ancak ikinci kez sorulduğunda verilmiş olmasıdır; zira "guten Menschen" ifadesinin ilk sorulduğunda nesne olarak değerlendirilmesi ve sonrasında cümlede ek eylem yapısı oluşturup oluşturmadığı sorgulandığında doğru cevabın verilmesi, modelin ilk soruya ürettiği yanıtı yeniden gözden geçirme ve bağlamsal unsurları ilişkilendirme kapasitesini göstermesi açısından önemlidir. Dil modeline sorulan ikinci sorunun yönlendirici nitelikte bir etki yapmış olabileceği göz önünde bulundurulmakla birlikte bu durum, sistemin plastisite işlevinin küçük ölçekte gözlenebildiği bir örnek olarak değerlendirilebilir. Sözü edilen örnek, aynı zamanda kullanıcı etkileşiminin model çıktıları üzerindeki etkisi bağlamında da tartışılabilir.

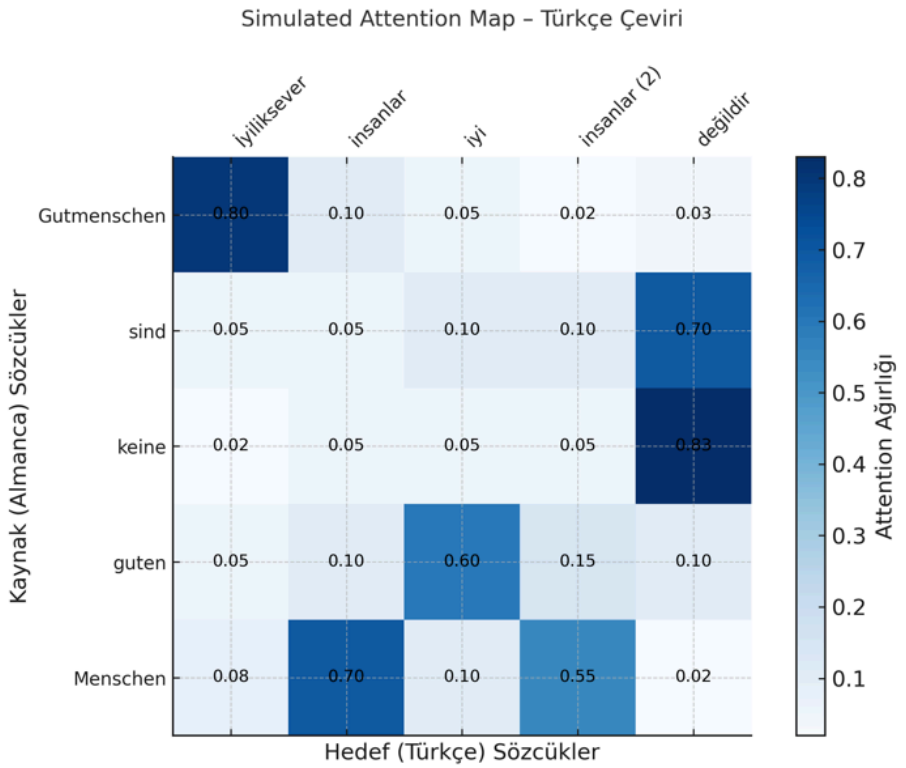
Örnek metin üzerinden devam edecek olursak *self-attention* mekanizması bu cümlede "Gutmenschen" ile "guten Menschen" ifadesi arasında olumsuzluk bildiren "keine" sözcüğü sayesinde güçlü bir karşıtlık ilişkisi kurmuştur ve "iyi görünmek" ve gerçekten "iyi olmak" arasındaki çelişkiyi tahmin etmiştir. "Gutmenschen sind keine guten Menschen" cümlesi, özellikle "Gutmenschen" sözcüğüne yapılan ironik veya eleştirel bir göndermedir. Burada "iyi görünme" çabası gerçek bir iyilikten ziyade dışsal bir gösteriş olarak ele alınır.

Yukarıdaki örnekler nöral makine çevirisinde sistemin özellikle İngilizce çevirisinde yalnızca sözcük düzeyinde bir eşleme yapmadığını; bu eşlemeyi, *attention* mekanizması aracılığıyla bağlamı tanıyarak ve anlamı yeniden kurarak gerçekleştirdiğini göstermektedir. *Self attention* mekanizması, modelin cümle içinde

özne-yüklem-karşıtlık ilişkisini fark etmesini sağlayarak bağlama duyarlı ve esnek bir biçimde çeviri çıktısı oluşturmasını sağlamıştır. Sistemin hangi sözcüklere ne düzeyde odaklanarak öğrendiği *backpropagation* ve *attention* mekanizmaları sayesinde anlaşılabilmektedir. Öğrenme süreci, sistemin daha önce karşılaştığı örneklerden edindiği esnek bilgiye dayanmakta; modelin sözcük düzeyinde karşılıktan ziyade bir metnin ironi, eleştiri tonu gibi farklı anlam katmanlarına hassas olduğunu ve sistemin plastisite kapasitesini devreye soktuğunu göstermektedir.

Görsel 3

Türkçe çeviri için dikkat haritası



Türkçe çeviriye gelecek olursak (bkz. Görsel 3) Türkçe çeviriye dair bir *attention map* oluşturulmuştur. ChatGPT arayüzü üzerinden gerçekleştirilen çeviri ve açıklamalar dikkate alınarak GPT-4 modeli tarafından simüle edilen dikkat haritası[9], sistemin hangi sözcükler arasında anlam kurduğunu ve hangi bağlamsal bileşenlere ne düzeyde dikkat verdiğini göstermektedir. Görsel, "Gutmenschen" sözcüğünün "İyiliksever" sözcüğü ile eşleştirildiğini, "sind" ve "keine" sözcüklerinin cümledeki olumsuz yargıyı vurguladığını ve bu doğrultuda "değildir" sözcüğüne yüksek ağırlık verdiğini göstermektedir. Kaynak metinde "Gutmenschen" sözcüğüne yapılan ironik veya eleştirel bir gönderme Türkçe çeviride çok fazla hissedilmemektedir. Modelin İngilizce çeviriye kıyasla bağlamsal tonlamayı aktarmadaki zayıflığı, dikkati çeken diğer bir noktadır. Bu durum, sistemin plastisite kapasitesinin zorlandığını ve yalnızca sözcük düzeyinde bir eşleştirme yaptığını göstermektedir. Bunun sebeplerine gelecek olursak, Türkçe-Almanca dil çifti arasındaki veri dengesizliği ve Türkçede tarz olarak bu tür gönderme, eleştiri ve ironi içeren söylemlerin veri içinde temsil edilme durumunun yetersizliği sayılabilir. Dil modellerinin plastisite kapasitenin artması için teknik olarak mimarilerin yanı sıra anlam çeşitliliği içeren bağlamsal farklılıklarla beslenmesi gerektiği söylenebilir. Ancak tekrar hatırlatmak gerekir ki seçilen metin, özgün bağlamı olan bir ifade içermekte ve çeviri anlamında büyük bir zorluk oluşturmaktadır.

[9]Dikkat haritası doğrudan modelin atadığı ağırlıkları gösterme noktasında temsilidir; zira GPT ya da diğer büyük dil modellerinde *attention* ağırlıklarına doğrudan erişim çoğu zaman mümkün değildir. Bunun sebebi ise decoder katmanlarının kapalı olmasıdır.

Metni bir bağlam içerisinde dil modeline verdiğimizde çeviride "Gutmensch" ifadesinin olduğu gibi aktarıldığı ve kaynak dildeki orijinal halini koruduğu görülmektedir. Metnin çevirisi için belli bir erek kitle ve amaç (*skopos*, vikipedi okuru, sözcük açıklaması) belirlediğimizde ise daha çok açıklayıcı çeviri yoluna gittiği görülmektedir:

Tablo 1

Kaynak-Erek Metin (https://de.wikipedia.org/wiki/Gutmensch)

Almanca Kaynak Metin	**Türkçe Erek Metin**
Gutmensch ist eine Bezeichnung, die häufig als ironisch oder verachtend gemeinte Verunglimpfung genutzt wird. Im Jahr 2015 wurde „Gutmensch" zum Unwort des Jahres gewählt. Es wird damit jemand bezeichnet, der naiv an das Gute glaubt. Ihnen wird aus Sicht der Wortverwender ein übertriebener Wunsch des „Gut-sein"-Wollens in Verbindung mit einem moralisierenden und missionierenden Verhalten und einer dogmatischen, andere Ansichten nicht zulassenden Vorstellung des Guten unterstellt.	Gutmensch, Almanca'da sıkça alaycı ya da küçümseyici bir anlamda kullanılan bir ifadedir. Genellikle bir kişiyi ya da grubu, naif bir iyilik inancına sahip olmakla ve aşırı ahlakçılıkla suçlamak için kullanılır. Bu terimle, kullanan kişiler açısından bakıldığında, "iyi insan olma" arzusunun abartılı biçimde dışa vurulması, başkalarını yargılayan ve kendi değerlerini dayatan bir tutumla ilişkilendirilir. Bu kişiler, çoğunlukla başka görüşlere kapalı, dogmatik bir iyilik anlayışıyla tanımlanır. "Gutmensch" kelimesi 2015 yılında Almanya'da "Yılın En Uygunsuz Sözcüğü" (Unwort des Jahres) seçilmiştir, çünkü farklı düşünenleri küçümsemek ve itibarsızlaştırmak amacıyla saldırgan bir biçimde kullanıldığı değerlendirilmiştir.

Yukarıdaki tablo bağlamın çeviri üzerindeki etkisini göstermek amacıyla oluşturulmuştur. Dil modeli, ideolojik bir vurgu taşıyan "Gutmensch" sözcüğüne tam bir Türkçe karşılık vermemiştir. Zira kültürel anlamda dile spesifik ifadelerin çevirisinde genellikle yabancı sözcüğün çevrilmeden bırakılması söz konusudur. Diğer taraftan metnin bir bağlam içerisinde verilmesi, sözcüğün taşıdığı kavramsal boyutun anlaşılmasına yardımcı olmuştur. Kaynak metinde geçen "Ihnen wird aus Sicht der Wortverwender ein übertriebener Wunsch des „Gut-sein"-Wollens in Verbindung mit einem moralisierenden und missionierenden Verhalten und einer dogmatischen, andere Ansichten nicht zulassenden Vorstellung des Guten unterstellt." cümlesinin erek dilde "Bu terimle, kullanan kişiler açısından bakıldığında, "iyi insan olma" arzusunun abartılı biçimde dışa vurulması, başkalarını yargılayan ve kendi değerlerini dayatan bir tutumla ilişkilendirilir." biçiminde bir cümle ile verildiği görülmektedir. Cümlede öznenin karışık bir yapıyla verilmesi ve yargı taşıyan sözcük olarak "ilişkilendirmek" eyleminin seçilmesi, özne yüklem arasında bir belirsizliğe sebep olmuştur. Böylelikle esneklikten uzak, komplike ve anlatım bozukluğu diyebileceğimiz bir yapı ortaya çıkmış, terim ile ne kast edildiği tam anlaşılmamıştır. Oysa metinde bu terim, onu kullananlar tarafından "iyi insan olma" arzusunun abartılı biçiminde dışa vurulmasıyla, başkalarını yargılayan, kendi doğrularını dayatan, farklı görüşlere kapalı bir tutumla özdeşleştirilmektedir. Metnin devamına kaynak metinde olmayan ancak bağlamın anlaşılmasına yardımcı cümleler eklendiği görülmektedir.

"Gutmensch" sözcüğünün Almanca siyasi bağlamda nasıl kullanıldığının yanı sıra toplumsal düzeyde işlevsel karşılığını erek okuyucuya fark ettirmeyi amaçlayan bir çeviri görevi belirlediğimizde ise GPT-4 (ChatGPT arayüzü üzerinden), *skopos'a* (Reiß ve Vermeer, 1984) uygun "iyilik budalası", "iyilik havarisi", "sözde iyi insan", "gösterişçi iyiliksever" gibi Türkçe işlevsel karşılıklar önermiştir.[10] Bu durum, dil modelinin bağlama göre farklı çeviri çıktıları oluşturduğunu ve bir noktaya kadar görece *skopos* doğrultusunda yönlendirilebildiğini göstermektedir. Bu bağlamda denilebilir ki plastisite, teknik bir yeterliliğin yanı sıra bağlama duyarlılık gerektirmektedir. Diğer taraftan genel anlamda bakıldığında nöral makine çevirisinde

[10]Siyasi bir tartışmada kullanıldığında ne kadar işlevsel olduğu sorgulanabilir. Türkçede buna benzer kullanımlara sosyal medyada çokça rastlanmıştır ancak "Gutmensch" kadar yerleşik Türkçe bir kullanım mevcut değildir. Genelde kavramın kendisinin Almanca olarak kullanıldığı örnekler yadsınamayacak ölçüdedir.

bu tarz metinlerin çözümlenmesi, kültürel ve ideolojik tonun korunması oldukça güç görünmekte, sistemin plastisitesinin zorlandığı görülmektedir.

Benzer biçimde Aslan'ın (2024) edebi metinlerde yapay zekâ destekli çeviri araçlarının performansını değerlendirdiği ve insan çevirmen ile karşılaştırdığı çalışması ile Odacıoğlu'nun (2024) hukuk metinlerinde yapay zekâ çevirisi ile insan çevirmeni karşılaştırdığı çalışması, yapay zekâ araçlarının bağlamsal duyarlılık ve plastisite bakımından kısıtlılıkları olabileceğini göstermektedir. İnsan çevirmen bu tarz çeviri görevlerinde öne çıkmaktadır. Bu noktada insan çevirmenin bağlamsal duyarlılık gerektiren özel bir çeviri görevinde art alan bilgisi, kültürel sermayesi[11], toplumsal deneyimini edimine nasıl entegre edeceği önemlidir. İnsan çevirmen, dilin nüansları ile deyimsel ve kültürel göndermelere hâkimiyeti sayesinde (Çetin ve Duran, 2024, s. 162) edinçlerini kullanarak, esnek ve bağlama uygun bir performans gösterebilir. Yapay zekâ ile çeviri kısıtlı verilerle sınırlıyken, insan kapasitesi itibariyle sonsuz bir evren gibidir (Şahin, 2023, s.9). Bu durum, işlevsel çeviri bakış açısından Nord'un (2005, s. 22-23) çevirmene uygun gördüğü iletişim tasarımcısı rolüyle bağdaşmaktadır. Zira günümüzde çevirmen, gerçekleştirilen iletişimi yöneten kişi konumuyla bütün sürecin sorumluluğunu taşımaktadır. Çeviri hem statik hem de bağlama göre şekillenen dinamik bir süreç (Hatim ve Mason, 2005, s. 24) olduğuna göre çevirmen, metni oluşturma sürecinde kaynak metin verilerinden yola çıkarak erek dil metninde hassas bir süreç yürütmeli; dilsel ve bağlamsal uyumu ustaca bir dikkatle gözetmeli ve erek dilde işlevsel bir ürün oluşturmalıdır. Nöral makine çevirisinde bu tarz düzenlemelerin karşılığı, *attention* mekanizmasıyla da bağlantılı olan *fine-tuning* uygulamasıdır. İnce ayar anlamındaki *fine-tuning* önceden eğitilmiş bir modelin belirli bir görev için yeniden eğitilmesidir. Yeniden eğitim sürecinde sistemin dikkat mekanizmaları optimize edilmektedir.

Fine-tuning- İnce ayar

Yapay zekânın plastisitesini sağlamayı hedefleyen mekanizmalardan biri de büyük dil modellerinde ön eğitim (*Pre-trained Models*) ve ince ayardır (*Fine tuning*) (Liang ve ark., 2021, s. 13333). Bazı dil modelleri (*BERT, GPT-3* vb.) çok büyük veri setleriyle önceden eğitilir ve ardından ince ayar yapılarak yeni bilgiye hızlı adapte olma özelliği ve esneklik kazanır.[12] Sözgelimi teknik, hukuk, tıp gibi alanlarda ek veri seti desteğiyle modelin son katmanlarına *fine-tuning* yapılabilir ve böylelikle modelin ağırlıkları (parametreleri) yeni veri setleriyle uyumlu olacak biçimde optimize edilir.

Esasında *fine-tuning*, nöral makine çevirisinde modelin önceki bilgilerini yeni durumlara adapte etmesini anlatan "transfer öğrenme" (*transfer learning*) paradigmasının bir parçası sayılabilir (Vrbančič ve Podgorelec, 2020). Transfer öğrenme sırasında, bir model büyük bir veri kümesi üzerinde eğitilir ve genel dil bilgisi, nesne tanıma veya başka geniş bilgi yapıları öğrenir. Tıpkı beynin daha önce öğrenilen bilgileri yeni görevlerde kullanabilmesi gibi, transfer öğrenme de bu bilgileri yeniden kullanarak yeni görevlerde daha hızlı öğrenmeyi sağlar (Kocmi, 2020). İnsan beyninin çalışma prensibi çok daha karmaşık işler; öğrenme süreçlerine duyusal, duygusal, sosyal ve çevresel faktörler eşlik eder. Nöral makine çevirisinde dil modelinin *fine-tuning* yoluyla belirli bir alanda teknik anlamda düzenlenmiş yeni bir bağlama adapte olması, daha önce bahsini ettiğimiz insan beyninde deneyime bağlı plastisitenin (*experience dependent plasticity*) yapay bir benzeri gibi görülebilir.

Bazı yapay zekâ mimarileri beyin benzeri bir esneklik sağlamak için dil modellerinde *fine-tuning'in* gelişmiş bir versiyonu gibi düşünülebilecek olan sürekli öğrenme (*continual learning*) yöntemi kullanmaktadır.

[11] Çevirmenin kültürel sermayesi, dilsel yetkinliğinin yanı sıra birikimi, kaynak ve erek kültür arasındaki ilişkisel ve sezgisel farkındalığı anlamında kullanılmıştır. Sermaye kavramı için bkz. Pierre Bourdieu, *Genel Sosyoloji: Collège de France Dersleri* (1981-1983) (Z. Emirosmanoğlu, Çev.), 2023.

[12] Kimi zaman da ince ayar sonucunda dil modeli yeni bilgiye aşırı uyum gösterir ve önceki bilgileri unutmaya eğilimli olur (*overfitting*- Liang ve ark., 2021).

Sözgelimi Elastik Ağırlık Konsolidasyonu (*EWC-Elastic Weight Consolidation*) gibi uygulamalar (Thompson ve ark., 2019; Variš ve Bojar, 2019), bu konuda kullanılan tekniklerdendir. Bu uygulama ile model, eski görevlerdeki ağırlıkları/parametreleri koruyarak yeni görevler için kendini optimize eder. Bunun amacı, modelin önceki bilgileri unutmamasıdır. Zira dil modellerinde bazen daha önce sözü edilen katastrofik unutma (*CF-catastrophic forgetting*)[13] sorunu yaşanmaktadır (Gu ve Feng, 2020; Wu, Liu ve Zong, 2024). Dil modelleri eğitim ya da ince ayar sırasında önceki bilgilerini unutabilmektedir (French, 2003). Bir örnekle somutlaştıracak olursak diyelim ki bir dil modeli Almanca-İngilizce dil çiftinde çeviri yapmak üzere eğitilmiş olsun ve performansı oldukça iyi olsun. Bu model, Almanca-Fransızca dil çiftine yönelik bir ince ayarla yeni bir göreve atanabilir. Model, normal bir *fine-tuning*'e tabi tutulduğunda Almanca-İngilizce dil çiftine yönelik parametreler güncellendiğinden modelin önceki parametreleri değişebilir ve çeviri performansı düşebilir. EWC bu noktada devreye giderek modelin işe yarayacak parametreleri korumasını sağlar ve modelin unutmasını engellemeye çalışır. Model böylelikle hem Almanca-İngilizce hem de Almanca Fransızca dil çiftinde performansını koruyarak başarılı çeviri yapabilir. Bu ve benzeri uygulamalardaki (*LwF- Learning without Forgetting, memory based approaches* vb.) amaç, modelin esnekliğini güncel ve işler tutmak, çeviri çıktılarının doğruluğunu ve güvenirliğini artırmaktadır.

Nöral makine çevirisinde bir dil modelinin belli bir alan ya da görev için uyarlanması amacıyla *fine-tuning* işlevine benzeri bir diğer yöntem ise geri getirme katmanı (*retrieval*) biçiminde dil modeline harici olarak eklenen *Retrieval-Augmented Generation* (*RAG*) uygulamasıdır (Gao ve ark., 2023). RAG yönteminde bilgi, *fine-tuning*'in aksine modele kalıcı olarak gömülmez, başka bir bilgi kaynağından ya da veri tabanından anlık sağlanır ve bilgi üretimine entegre edilir (*non-parametric*). Böylece model, kendi parametrelerinde kayıtlı olan bilginin yanı sıra güncel ve özel verilere odaklanarak bağlamsal esnekliği ve doğruluğu artırmaya çalışır. RAG özellikle teknik bilgiye sahip olmayan kullanıcıların, ek yazılım geliştirme veya model eğitimi yapmadan alan odaklı ve güncel verilere erişmesine olanak tanıyabilir. Buna karşılık *fine-tuning* ile uyarlanmış bir dil modeli eğitildiği veri dışındaki güncel bilgilere erişemez ancak alan/özel bilgileri modelin parametrelerinde kalıcı olduğu için (*parametric*) uzun vadeli uyarlama gerektiren durumlar için avantaj sağlayabilir.

Uygulama Örneği: MultiMaCoCu

Bu bölümde çalışmamızda ele alınan plastisite olgusunun nöral makine çevirisi bağlamında nasıl somutlaştığını göstermek amacıyla bir *fine-tuning* uygulamasına yer verilmektedir. Uygulama, Türkçe-İngilizce çift dilli veri üzerinde çalışan, *Transformer* mimarisi tabanlı dil modeli *MultiMaCoCu* kullanılarak gerçekleştirilmiştir. Bu bağlamda amaç, modelin alan/özel veri ile yeniden eğitilmesi sürecini, elde edilen çıktıları ve bu süreçte gözlemlenen bağlamsal esnekliği ortaya koyarak plastisite kavramının çeviri teknolojilerindeki yansımasını açıklamaktır.

Uygulama Adımları

Dil modelinin eğitimi, çeviri ve dil modelleri alanında İstanbul Üniversitesi'nde doktora çalışmalarını yürüten Burak Özkaracahisar'ın teknik altyapının kurulumu, veri ön işleme süreci ve model parametrelerinin optimize edilmesine yönelik katkılarıyla uygulanmıştır.[14] Tercih edilen modelin temel eğitimi, OpenNMT-py altyapısı üzerinden geniş çaplı çok dilli veriyle gerçekleştirilmiştir. Model, bu yönüyle belirli bir alanın terminolojisine ve bağlamına uyum sağlama potansiyeliyle öne çıkmaktadır.

[13] Catastrophic forgetting, catastrophic interference ile aynı anlamdadır.

[14] Dil modeli eğitimi ve *fine-tuning* uygulanması sürecindeki katkıları için Burak Özkaracahisar'a teşekkürü borç bilirim. Modelin eğitimi için kullanılan veri setlerinin alındığı kaynaklar ise şöyledir: https://opus.nlpl.eu/MultiMaCoCu/en&tr/v2/MultiMaCoCu https://opus.nlpl.eu/ELRC-3057-wikipedia_health/en&tr/v1/ELRC-3057-wikipedia_health

Veri hazırlığı ve ön işleme aşamasında Türkçe ve İngilizce dillerinde sağlık alanına ait makale, rapor ve kılavuzlardan oluşan çift dilli bir derlem toplanmıştır. Toplanan veri seti hem kaynak hem de erek dilde temizlenerek biçimsel hatalar giderilmiştir. *Fine-tuning* süreci öncesinde tüm veriler *subword tokenizasyon* adı verilen alt birimlemeye tabi tutulmuştur. Verilerin alt birimlere ayrılmasının ve bu şekilde işlenmesinin amacı, modelin sözcükleri anlamlı alt parçalar halinde temsil etmesini sağlayarak alana özel terimlerin tutarlı öğrenilmesini sağlamaktır.

Fine-tuning süreci, OpenNMT-py altyapısı kullanılarak oluşturulan çift dilli veri seti üzerinde yürütülmüştür. *Fine-tuning* aşamasında sağlık metinleri, genel metinlere oranla 10 kat daha fazla ağırlıklandırılmış, modelin sağlık alanına yönelimi artırılarak yüksek duyarlı olması hedeflenmiştir. Eğitim süreci 5000 adım olarak planlanmış, her 1000 adımda bir kontrol noktası (*checkpoint*) kaydedilmiştir. Daha önce sözünü ettiğimiz aşırı öğrenmeyi (*overfitting*) önlemek amacıyla değerlendirme sonuçları doğrultusunda 3000 adım son model olarak seçilmiştir. Ve son olarak çeviri metinleri yine alt sözcüklere ayrıldıktan sonra model, çeviri çıktısı üretmek amacıyla kullanılmıştır. Üretilen çeviriler, önceden eğitilmiş temel modelin çıktılarıyla karşılaştırılmıştır.

MultiMaCoCu Çeviri Çıktıları

Bu bölümde *MultiMaCoCu* dil modelinden *fine-tuning* uygulama sonrası elde edilen çeviriler, temel model çevirileriyle karşılaştırılmış, modelin alana uyum duyarlılığı incelenmiştir (bkz. Tablo-2). Ardından dil modeline çevrilmek üzere haricen bir makale verilmiş ve çeviri çıktıları incelenmiştir (bkz. Tablo-3).

Tablo 2

MultiMaCoCu Dil Modeli Çeviri Çıktıları

Kaynak Metin İngilizce	Temel Model Çevirisi	Fine-tuning Uygulanmış Çeviri
While most *cases* are due to bleeding from small aneurysms, larger aneurysms (*which are less common*) are more likely to *rupture*.	Çoğu *olguda* ufak anevrizmalardan kanama olsa da daha büyük anevrizmalar (*daha az yaygın olan*) *yırtılma* olasılığı daha yüksektir.	Çoğu *vaka* küçük anevrizmaların kanamasından kaynaklansa da, daha büyük anevrizmaların (*daha nadir*) *rüptüre* olma olasılığı daha yüksektir.

Çeviri çıktılarında görüldüğü üzere İngilizce kaynak metinde geçen "case" sözcüğü temel model çevirisinde "olgu" olarak; *fine-tuning* uygulanmış modelde ise "vaka" olarak karşılanmıştır. Tıbbi metinlerde her iki kullanıma rastlanmakla birlikte, modelin farklı sözcük tercihlerinin, eğitim sürecindeki eşleşmelerin istatistiksel ağırlığına göre biçimlendiği düşünülebilir. Bunun yanı sıra temel model "which are less common" ifadesini "daha az yaygın olan" biçiminde çevirirken, *fine-tuning* uygulanmış model tıptaki yaygın kullanımına uygun "daha nadir" ifadesini seçmiştir. Temel model ile *fine-tuning* uygulanmış model çevirileri arasındaki en belirgin fark ise kaynak metinde geçen "rupture" sözcüğünde görülmektedir. Sözcüğün çevirisinde temel model genel dile ait "yırtılma" sözcüğünü seçerken, alan uyarlaması yapılmış metinde tıbbi terim olan "rüptür/rüptüre olma" ifadesinin kullanıldığı görülmektedir. Bu gözlemler, *fine-tuning* sayesinde modelin terminolojik duyarlılığının arttığını göstermektedir.

Aşağıda Herloz ve arkadaşlarının çalışmasından (2012, s.3) modele verilen makale metnine ait çeviri çıktıları yer almaktadır. Beyin tümörlerinde tanı ve tedavi yöntemlerinin ele alındığı makalede PET (*Pozitron Emisyon Tomografisi*) adı verilen bir görüntüleme yönteminden söz edilmektedir. Görüntüleme işleminde FDG (*Fluorodeoksiglukoz*) adı verilen bir madde kullanılmaktadır. FDG, glukoza benzeyen bir molekül olduğu için vücuda enjekte edildiğinde hücreler tarafından glukoz gibi alınır ve PET taramasında bu maddenin tutulduğu bölümler daha parlak görünür. Görüntülemede FDG kullanılmasının amacı, tümörün olduğu bölümlerin tespit edilmesi, sınırlarının belirlenmesidir. Makalede konuya ilişkin bir bölüm Türkçeye çevrildiğinde temel model ve *fine-tuning* uygulanmış modelin çeviri çıktıları tabloda sunulmuştur:

Tablo 3
MultiMaCoCu Makale Çevirisi Çıktıları

Kaynak Metin İngilizce	Temel Model Çevirisi	Fine-tuning Uygulanmış Çeviri
The high and regionally variable FDG *uptake* in normal brain often makes the *delineation* of brain tumors difficult.	Normal beyinde yüksek ve bölgesel olarak değişken FDG *alımı* genellikle beyin tümörlerinin *gidişini* zorlaştırmaktadır.	Normal beyindeki yüksek ve bölgesel olarak değişken FDG *alınımı*, sıklıkla beyin tümörlerinin *delineasyonunu* zorlaştırmaktadır.
Quantitation of FDG *uptake* in brain and tumors....	Beyinde ve tümörlerde FDG *tutulum* miktarı....	Beyinde ve tümörlerde FDG *tüketim* miktarı....

Yukarıda verilen çeviri çıktılarında görüldüğü üzere cümlede geçen "tanımlama, sınır belirleme" anlamında kullanılan "delineation" sözcüğü temel model çevirisinde "tümörün gidişi" biçiminde karşılık bulurken, *fine-tuning* uygulanmış modelde "delineasyon" biçiminde ödünçleme bir sözcük olarak karşımıza çıkmaktadır. Temel modelin çeviri çıktısında sözcük tercihiyle cümlenin anlamı muğlaklaşmaktadır; buna karşın *fine-tuning* uygulanmış çeviri çıktısında sağlık alanında terminolojiye duyarlılığın yüksek olduğu ve bu doğrultuda teknik bir terim tercih edilerek modelin esnekliğinin sağlandığı gözlemlenmektedir. Öte yandan başka bir cümlede FDG'nin hücre tarafından alınmasını/emilimini anlatan "uptake" sözcüğünün temel modelde "tutulum" ve diğer modelde "tüketim" olarak çevrildiği görülmektedir. Yüzeysel baktığımızda bu tercihin terminolojik bir uyum sağlamak adına gerçekleştiği düşünülse de FDG'nin metabolik olarak tüketilmediği, hücre tarafından tutulduğu dikkate alındığında *fine-tuning* uygulanmış modelde çevirinin anlamı olumsuz yönlendirme potansiyeli taşıdığı görülmektedir. Bu tarz sınırlılıkların iyileştirilmesi için, modelin üzerinde çalışılması[15], geri bildirim mekanizmalarıyla desteklenmesi, alan/özel paralel çeviri çiftleriyle (*parallel corpora*) zenginleştirilmesi gibi işlemlerin uygulanması gerektiğini belirtmek gerekir. Zira *fine-tuning* uyguladığımız modelin nihai bir ürün olmadığını; plastisiteyi incelemek amacıyla deneysel bir uygulama olarak tasarlandığını bu noktada hatırlatmak önemlidir.

Bulgular ve Tartışma

Çalışmada nöral makine çevirisinde plastisiteyi sağlayan temel mekanizimaların çeviri çıktıları üzerindeki etkisi analiz edilmiştir. *Backpropagation* mekanizması altında verilen örnekler, nöral makine çevirisinde plastisitenin, dilsel çözümlemenin ötesinde sistemin beslendiği veri kaynaklarıyla sağlandığını göstermektedir. Bunda kullanılan mimarinin ve verimli öğrenmenin de etkisi vardır. *Attention* mekanizması ve dikkat haritaları ile incelenen örnekler, modelin kavram çiftleri arasında ilişki kurabildiğini ve yapısal düzeyde uyum sağladığını gösterirken, pragmatik düzeyde plastisitenin yetersiz kaldığı görülmüştür. Son olarak *MultiMaCoCu* adlı modelden tasarlanan ve eğitilen dil modeline uygulanan *fine-tuning* sonucunda model, eğitim öncesine göre bağlamsal duyarlılığı belli ölçüde sağlamıştır ancak plastisite konusunda bazı kısıtlılıklar saptanmıştır. Bulgular, nöral makine çevirisinin plastisiteye dayalı verimlilik potansiyelinin geliştirilmeye açık olduğunu ancak bununla birlikte modelin işleyişinde istenmeyen yan etkilere de yol açabileceğini göstermiştir.

Plastisite, dil modellerinde ve nöral makine çevirisinde bağlamsal esnekliği sağlayan bir yeti olarak tanımlansa da sözü edilen esnekliğin ne ölçüde ve nasıl gerçekleştiğinin tartışılması gerekmektedir. *Backpropagation* mekanizması dil modelinin hatalarını esas alarak modelin yeniden yapılanması ile karakterizedir. Ne var ki *backpropagation*, insan beyni ile kıyaslandığında Hinton'un ifade ettiği gibi teknik ve örüntüye dayalı bir düzeltme mekanizmasıdır. Bu durum, çeviri söz konusu olduğunda bağlamın yalnızca

[15]Genel cümleler yerine alan terimlerine, mecazi ifadelere, kısaltmalara farklı ağırlıklar verilebilir. Sözgelimi genel cümleler 1, alan terimleri 8, mecazi anlatımlar 2, kısaltmalar 10 biçiminde model parametreleri optimize edilerek *fine-tuning* gerçekleştirilebilir.

örüntüyü izlemesi ile sonuçlanır. İnsan beyninde olduğu gibi çeşitli bilişsel süreçleri takip etmediği için *backpropagation* mekanizmasının plastisiteyi sağlamak konusunda kısıtlılıkları olduğu açıktır. Nöral sinir ağlarında sistemin genel işleyişine entegre olan *attention* mekanizması ise dil modelinin metni işlerken sözcükler arasında ilişki kurmasını sağlayan, bağlama göre bağlantısallığı öğrenmeye çalışan önemli bir plastisite aracıdır. Metin üzerinden yapılan analizde görüldüğü üzere dikkat mekanizması, sözcük düzeyinde ve sözdizimsel düzeyde bir eşleşme gerçekleştirirken söylemsel düzeyde plastisite sağlamak bakımından sınırlı kalabilmektedir. *Fine-tuning*'e gelince dil modelinin yeni bir duruma ve bağlama uyumlanabilirliğinin sınanması söz konusudur. Burada plastisitenin teknik anlamda bir örüntü takip etmenin ötesinde, modelin beslendiği veri ile ilgili olduğu ortaya çıkmaktadır.

Yapılan incelemeler yapay zekânın plastisitesiyle insan çevirmenin plastisitesinin farklı işleyiş mekanizmalarına sahip olduğunu göstermektedir. İnsan çevirmen, donanımıyla dilin yanı sıra birçok farklı etkeni sürece dahil ederek bilişsel sürecini aktif kılmasıyla, çeviri sürecinde kendisi dışında diğer etkenlerle etkileşimi ve süreci yönetmesiyle Nord'un belirttiği üzere tam bir metin tasarımcısıdır. Yapay sistemler bu sürecin önemli bir parçası olabilir ancak insan, henüz başat bir rol oynamaya devam etmektedir. Bu rolün ne kadar süreceği ya da neye dönüşeceği farklı gelişmelere göre biçimlenecektir. Bunun yanı sıra yapay zekânın ve insan çevirmenin süreçteki görevleri, iş birlikleri ile ayrıca her iki aktörün avantajları ve dezavantajları tartışılabilir. Zira çeviribilim ile teknoloji arasındaki ilişki sürekli kendini yenileyen bir niteliktedir (Genç ve Çınar, 2024, s. 131).

Öte yandan plastisite başlığı altında çevirmenin yanlılığı ile yapay zekânın karar verme süreçlerindeki girdi-çıktı ilişkisi üzerine düşünmek gerekir. Yapay zekâ araçlarıyla etkileşimde çeviri görevinin nasıl tanımlandığı; metnin çözümlenme aşaması ve bu doğrultuda çeviri yönergesinin nasıl oluşturulduğu; üslup, işlev, terminoloji, metin uzunluğu, tercihler ve yasaklarla ilgili hangi kısıtların prompa sunulduğu, verilen bağlamın kapsamı gibi unsurlar modelin çıktı üretirken dikkatini ve odağını (*attention*) yeniden düzenler, modelin olası seçimlerini genişletir, daraltır ya da önceliklendirir. Böylelikle dil modelinin plastisitesinin bağlamsal uyum bakımından nasıl biçimleneceği gözlenebilir. Anlaşılan o ki plastisite teknik bir altyapıya dayanmakla birlikte kullanıcı etkileşimi ile de yön değiştirebilen bir niteliğe sahiptir. Çalışmada "Gutmensch" örneği bunu açıkça göstermektedir. Bağlam verilmediğinde modelin "iyi insan" biçiminde nötr bir çıktı üretmesi karşılığında "siyasal bağlam ve aşağılayıcı işlev" gibi bağlamı tanımlayıcı yönergelerin etkisiyle "iyilik budalası", "iyilik havarisi" gibi farklı olasılıklar üretmesi, kullanıcı-model etkileşiminin çeviri kararlarını nasıl etkilediğini ortaya koymaktadır. Bu noktada Kumlu ve Okul'un (2023) insan çevirisi ile GPT-3.5 çıktılarını karşılaştırdıkları çalışmadan söz edilebilir. Çalışmada çeviri örneklerinin "bağlamın anlaşılması için yeterli uzunlukta" (s.34) seçilmesi ve modele "Bu metni bir edebiyat çevirmeni olduğunu hayal ederek ve kültürel unsurlara dikkat ederek İngilizce diline çevirebilir misin?" biçiminde bir komut verilmesi neticesinde modelin ürettiği çıktı ve ona eşlik dönütler, model davranışının bağlam girdisine duyarlı olduğunu görünür kılmaktadır.

Bu doğrultuda çevirmenin çeviri çıktısı üretmek dışında farklı roller üstlenebileceği günlerin çok uzak olmadığı söylenebilir. Çevirmen, yapay zekâ sistemlerini verimli kullanmayı öğrenerek birçok etkin görev alabilir. Bunun için öncelikle yapay zekânın işleyişi ve mekanizmaları ile ilgili bir kavrayış geliştirmelidir. Böyle bir zemin çevirmeni, süreci tasarlayan ve denetleyen çok işlevli bir aktör olarak komunlandırabilir. Sektörde çevirmenlerin en öne çıkan katkıları, üretim öncesi ön düzenleme (*pre-editing*) ile üretim sonrası son düzenleme (*post-editing*) işlevlerinde toplanmaktadır (Şahin, 2023, s.111). Ön düzenleme, metin ve bağlamın yapay zekâya sunulmadan önce tasarlanmasıdır. Bağlamın tasarlanması, çevirinin işlevi, amacı, temsil gücü yüksek örnek metinleri, gerekli arka plan bilgisini içeren bir girdi dosyası oluşturmayı ve bu girdi dosyasına göre verilecek komutları belirlemeyi gerektirebilir. Bu adım, modelin odaklarını ve seçim olasılıklarını baştan düzenleyerek çıktı kalitesini belirleyici biçimde etkileyebilir. Söz konusu görevler, bağlam tasarımcısı, bağlam editörü, prompt uzmanı gibi başlıklar altında uzmanlaşabilir. Son düzenlemede ise metnin kontrolü,

test edilmesi için gerekli metin havuzu oluşturmak, telif, etik gibi durumları denetlemek gibi farklı roller ortaya çıkabilir. Bunların yanı sıra yapay zekâ ile çeviride en önemli hazine "veri" dir (Şahin, 2023, s. 113). Çevirmen, veri setleri ile çalışarak girdi-çıktı arasındaki ilişkiyi düzenleyen ve güçlendiren derlem yöneticisi görevi üstlenebilir.

Sonuç

Çalışmada *backpropagation* ve *fine-tuning* gibi teknik anlamda esnekliği sağlamaya yardımcı mekanizmalar ile sisteme entegre olarak işlev gören *attention* bileşeninin metinler üzerinden incelenmesi neticesinde bu mekanizmaların görece plastisiteyi desteklediği ancak bazı kısıtlılıkları olduğu görülmüştür. Sonuç olarak plastisite, mevcut sistemlerin bağlamsal çeşitliliğe duyarlılık geliştirme kapasitesiyle ilgili bir konudur. Plastisitenin gelişmesinde diğer bir faktör, mevcut sistemlerin beslendiği veri kaynaklarının belli bir çeşitliliğe ve zenginliğe ulaşmasıdır. Bunun yanı sıra kullanıcı-model etkileşimi, plastisite bakımından belirli düzeyde yönlendirici bir işlev görmektedir. Dolayısıyla elde edilen bulgular, çeviri ediminin yalnızca teknik bir aktarım değil, bağlamsal ve kültürel boyutlarıyla bütüncül ele alınması gereken bir süreç olduğunu göstermektedir. Bu nedenle, insan ve nöral makine çevirisinin güçlü yönlerinin birleştiren çevirmen-yapay zekâ etkileşimine odaklanmak, çeviri araştırmaları bakımından kaçınılmaz hale gelmektedir.

Son Notlar

İncelenen ve analiz edilen örnekler, nöral makine çevirisinin kapasitesine dair temsili bir nitelik taşımaktadır. Bu doğrultuda plastisite ile ilgili yapılacak çalışmaların farklı dil çiftleri, büyük veri setleri ve farklı kullanıcı deneyimleri ile derinleştirilmesi önerilmektedir. Bunun yanı sıra plastisitenin dilde hangi unsurlarla sınanabileceği konusunda kültürel ve ideolojik söylemlere ilişkin bir veri evreni oluşturulabilir ve çeviride işlevsel karşılıklarla ilgili çeşitli dil modelleri eğitilerek çalışmalar zenginleştirilebilir.

Teşekkür İstanbul Üniversitesi Çeviribilim Bölümü'nde doktora yapan ve dil modeli eğitiminde bana yardımcı olan Burak Özkaracahisar'a çok teşekkür ederim.
Hakem Değerlendirmesi Dış bağımsız.
Çıkar Çatışması Yazar çıkar çatışması bildirmemiştir.
Finansal Destek Yazar bu çalışma için finansal destek almadığını beyan etmiştir.

Acknowledgment I would like to thank Burak Özkaracahisar, who is a PhD student at the Department of Translation Studies at Istanbul University and helped me with the language model training.
Peer Review Externally peer-reviewed.
Conflict of Interest The author has no conflict of interest to declare.
Grant Support The author declared that this study has received no financial support.

Yazar Bilgileri / Author Details
Derya Oğuz (Doç.Dr.)
[1] Marmara Üniversitesi, İnsan ve Toplum Bilimleri Fakültesi, Mütercim ve Tercümanlık Bölümü, İstanbul, Türkiye
0000-0001-5228-076X derya.oguz@marmara.edu.tr

Kaynakça | References

Abbas, Z., Zhao, R., Modayil, J., White, A. & Machado, M. C. (2023). Loss of plasticity in continual deep reinforcement learning. In *Conference on lifelong learning agents*, 232: 620-636. PMLR. Retrieved from: https://proceedings.mlr.press/v232/abbas23a.html

Açıkgöz, N. & Madi, B. (2013). Öğrenme ile Beyinde Oluşan Değişiklikler (Plastisite). *Marmara Üniversitesi Atatürk Eğitim Fakültesi Eğitim Bilimleri Dergisi*, 9(9), 29-36.

Allemann, A., Atrio, À. R. & Popescu-Belis, A. (2024). Optimizing the Training Schedule of Multilingual NMT using Reinforcement Learning. Preprint *arXiv*, Computation and Language. https://doi.org/10.48550/arXiv.2410.06118

Anlar, B. (2013). Beyinde Plastisite ve Bozuklukları. *Türkiye Klinikleri Pediatrik Bilimler- Özel Konular*, 9(4), 129-137.

Aslan, E. (2024). Yapay Zekâ Destekli Çeviri Araçlarının Edebi Çevirideki Yeterlilikleri Üzerine Karşılaştırmalı Bir İnceleme. *IU Journal of Translation Studies*, (20), 32-45.

Bahdanau, D., Cho, KH. & Bengio, Y. (2015). Neural machine translation by jointly learning to align and translate. In *Proceedings of ICLR*. Preprint *arXiv*, https://doi.org/10.48550/arXiv.1409.0473

Baker, M. (1992). *In Other Words: A Coursebook on Translation*. Routledge, London and New York. Retrieved from: https://archive.org/details/InOtherWordsByMonaBaker/page/n3/mode/2up

Bourdieu, P. (2023). *Genel Sosyoloji: Collège de France Dersleri (1981-1983)* (Z. Emirosmanoğlu, Çev.), İletişim Yayınları.

Chen, N. (2024). Text Classification Model Based on Long Short-Term Memory with L2 Regularization. In *2024 Second International Conference on Data Science and Information System (ICDSIS)*, 1-4. IEEE. Doi: 10.1109/ICDSIS61070.2024.10594621

Retrieved from: https://www.researchgate.net/publication/382377464_Text_Classification_Model_Based_on_Long_Short-Term_Memory_with_L2_Regularization

Corbetta, M. & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature reviews neuroscience*, 3(3), 201-215. Doi: 10.1038/nrn755 https://www.researchgate.net/publication/11375373_Control_of_Goal-Directed_and_Stimulus-Driven_Attention_in_the_Brain

Cüceloğlu, D. (1995). *İnsan ve Davranışı: Psikolojinin Temel Kavramları*, 25. Basım Remzi Kitapevi.

Çetin, Ö., & Duran, A. (2024). A Comparative Analysis Of The Performances of ChatGPT, DeepL, Google Translate And A Human Translator In Community Based Settings. *Amasya Üniversitesi Sosyal Bilimler Dergisi*, 9(15), 120-173.

DWDS *Digitales Wörterbuch der Deutschen Sprache*. (2025, 24 Mayıs). https://www.dwds.de/wb/Verfahren

Dohare, S., Fernando Hernandez-Garcia, J., Lan, Q., Rahman, P., Rupam Mahmood, A. & S.Sutton, R. (2024). Loss of plasticity in deep continual learning. *Nature*, Springer Science and Business Media LLC. https://doi.org/10.1038/s41586-024-07711-7

French, R. M. (2003). Catastrophic interference in connectionist networks. Nadel L. (Ed.) In *Encyclopedia of Cognitive Sciences* (Vol.1, s. 431-435). London: Nature Publishing Group. Retrieved from: http://leadserv.u-bourgogne.fr/rfrench/french/catastrophic_forgetting.ECS.french.pdf

Gao, Y., Xiong, Y., Gao, X., Jia, K., Pan, J., Bi, Y., ... & Wang, H. (2023). Retrieval-augmented generation for large language models: A survey. *arXiv preprint arXiv:2312.10997*, 2(1).

Genç, A. & Çınar Yağcı, Ş. (2024). Çeviribilim Alanında Yapay Zekâ Üzerine Ulusal Alan Yazında Yazılmış Makalelerin Eğilimleri Üzerine Bir Araştırma. *Kırıkkale Üniversitesi Sosyal Bilimler Dergisi*, 14(3), 119-136.

Gu, S. & Feng, Y. (2020). Investigating catastrophic forgetting during continual training for neural machine translation. In *Proceedings of the 28th International Conference on Computational Linguistics*, 4315–4326, Barcelona, Spain (Online). International Committee on Computational Linguistics. Doi: 10.18653/v1/2020.coling-main.381 Retrieved from: https://aclanthology.org/2020.coling-main.381.pdf

Hatim, B. & Mason, I. (2005). *The Translator as Communicator*. Taylor & Francis e-Library. First pub. (1997) Routledge, London and New York. ISBN 0-203-99272-5

Herholz, K., Langen, K. J., Schiepers, C. & Mountz, J. M. (2012). Brain tumors. In *Seminars in nuclear medicine* 42 (6), 356-370. WB Saunders. Doi:10.1053/j.semnuclmed.2012.06.001.

Hinke, R. M., Hu, X., Stillman, A. E., Kim, S.-g., Merkle, H., Salmi, R. & Ugurbil, K. (1993). Functional magnetic resonance imaging of Broca's area during internal speech. Neuroreport: *An International Journal for the Rapid Communication of Research in Neuroscience*, 4(6), 675–678. https://doi.org/10.1097/00001756-199306000-00018

Jääskeläinen, R. (2011). Focus on methodology in think-aloud studies on translating. In *Tapping and mapping the processes of translation and interpreting: Outlooks on empirical research* (s. 71-82). John Benjamins Publishing Company. https://doi.org/10.1075/btl.37.08jaa

Kerimoğlu, C. (2022). Dilin Kökeni Arayışları-5: Beyin ve Dil. *Dil Araştırmaları*, 16(30), 21-37. https://doi.org/10.54316/dilarastirmalari.1075944

Kılıç, T. (2024). *Yeni Bilim: Bağlantısallık Yeni Kültür: Yaşamdaşlık*, 7. Basım, Ayrıntı Yayınları, İstanbul. Birinci Basım 2021.

Kocmi, T. (2020). *Exploring Benefits of Transfer Learning in Neural Machine Translation.* (Doctoral dissertation, Institute of Formal and Applied Linguistics, Charles University). Retrieved from: https://dspace.cuni.cz/bitstream/handle/20.500.11956/115854/140081447.pdf?sequence=1

Kumlu, D., & Okul, M. (2023). Yapay Zekâ ile Çeviri Uygulamaları: Kültürel Unsurların Çevirisinde ChatGPT. *Contemporary Translation Studies*, 24.

Liang, J., Zhao, Ch., Wang, M., Qui, L. & Li, L. (2021). Finding Sparse Structures for Domain Specific Neural Machine Translation. *The Thirty-Fifth AAAI Conference on Artificial Intelligence (AAAI-21)*. Retrieved from: https://cdn.aaai.org/ojs/17574/17574-13-21068-1-2-20210518.pdf

Nord, C. (2014). *Translating as a Purposeful Activiy: Functionalist Approaches Explained*. Routledge, London and New York. First pub. (1997), St. Jerome Publishing.

Odacıoğlu, M. C. (2024). Yapay Zekâ ve İnsan Çevirisi: Hukuk Metinlerine Dayalı Karşılaştırmalı Bir Çalışma. *Karamanoğlu Mehmetbey Üniversitesi Uluslararası Filoloji ve Çeviribilim Dergisi*, 6(1), 147-171. https://doi.org/10.55036/ufced.1477008

PACTE GROUP. (2017). Decision-making. In *Researching Translation Competence by PACTE Group* (191-210). John Benjamins Publishing Company.

Posner, M. I. & Petersen, S. E. (1990). The attention system of the human brain. *Annual review of neuroscience*, 13(1), 25-42. Doi: 10.1146/annurev.ne.13.030190.000325 Retrieved from: https://www.researchgate.net/publication/20971732_The_Attention_System_of_the_Human_Brain

Price, C. J., Green, D. W. & Von Studnitz, R. (1999). A functional imaging study of translation and language switching. *Brain*, 122(12), 2221-2235.

Reiß, K. & Vermeer, H. (1984). *Grundlegung einer allgemeinen Translationstheorie*. De Gruyter.

Rumelhart, D. E., Hinton, G.E. & Williams, R.J. (1986). Learning representations by back-propagating errors. *Nature*, 323(6088), 533-536.

Şahin, M. (2023). *Yapay Çeviri*. Çeviribilim Yayınları.

Schlegel, P., Yin, Y., Bates, A. S., Dorkenwald, S., Eichler, K., Brooks, P., ... Jefferis, G. S. (2024). Whole-brain annotation and multi-connectome cell typing of Drosophila. *Nature*, 634(8032), 139-152. https://doi.org/10.1038/s41586-024-07686-5

Snelleman, E. (2016). *Decoding neural machine translation using gradient descent* (Master Thesis in Computer Science, Algorithms, Language and Logic, Department of Computer Science and Engineering, Chalmers University of Technology, Gothenburg). Retrieved from: https://odr.chalmers.se/server/api/core/bitstreams/c2b7d271-c5f9-472c-b0b7-36bd8eef1d40/content

Stanford Online (2016, 27 Nisan). *Can the brain do back-propagation? Geoffrey Hinton*. [Video]. YouTube. https://www.youtube.com/watch?v=VIRCybGgHts

Teke Tek Bilim (2024, 26 Mayıs). *İnsan beyninin sınırları? Prof. Dr. Türker Kılıç & Fatih Altaylı*. [Video]. YouTube. https://www.youtube.com/watch?v=jUZvlU1It5U

Thompson, B., Gwinnup, J., Khayrallah, H., Duh, K. & Koehn, P. (2019). Overcoming catastrophic forgetting during domain adaptation of neural machine translation. In *2019 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, 2062-2068. Retrieved from: https://aclanthology.org/N19-1209.pdf

Variš, D. & Bojar, O. (2019). Unsupervised pretraining for neural machine translation using elastic weight consolidation. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics: Student Research Workshop*, 130-135. Florence, Italy. Retrieved from: https://aclanthology.org/P19-2017.pdf

Vaswani, A., Shazeer, N., Parmar, N., Uskoreit, J., Jones, L., Gomez, A. N., Kaiser. L. & Polosukhin, I. (2017). Attention Is All You Need. In *31st Conference on Neural Information Processing Systems*, 5998-6008. Long Beach, CA. https://doi.org/10.48550/arXiv.1706.03762

Vig, J. (2019). A multiscale visualization of attention in the transformer model". In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, 37–42, Florence, Italy. Preprint *arXiv* https://doi.org/10.48550/arXiv.1906.05714

Vrbančič, G. & Podgorelec, V. (2020). Transfer learning with adaptive fine-tuning. *IEEE Access*, 8, 196197-196211.Retrieved from: https://www.researchgate.net/publication/345713638_Transfer_Learning_With_Adaptive_Fine-Tuning

Wenger, E. & Lövdén, M. (2016). The learning hippocampus: Education and experience-dependent plasticity. *Mind, Brain, and Education*, 10(3), 171-183. https://doi.org/10.1111/mbe.12112

Wu, J., Liu Y. & Zong C. (2024). F-MALLOC: Feed-forward Memory Allocation for Continual Learning in Neural Machine Translation. In *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 7180–7192, Mexico. Association for Computational Linguistics. Main conference of NAACL. Retrieved from: https://aclanthology.org/2024.naacl-long.398/

Zhao, Y., Zhang, J. & Zong, C. (2023). Transformer: A general framework from machine translation to others. *Machine Intelligence Research*, 20(4), 514-538. Doi: 10.1007/s11633-022-1393-5 Retrieved from: https://nlpr.ia.ac.cn/cip/ZongPublications/2023/2023-ZhaoYang-MIR.pdf