


Sözcüksel Muğlaklığın Doğal Dil İşleme Sistemlerindeki Yeri †

[The Role of Lexical Ambiguity in Natural Language Processing Systems]

Tolgahan TOY 

Middle East Technical University

Received: 30.11.2022 / Accepted: 09.05.2023

DOI: 10.51404/metazihin.1212525

Research Article

Abstract: What are the basic tenets of a system that understands natural languages? What obstacles need to be overcome? One of the problems related to natural language understanding processes is ambiguity. There are two types of ambiguity: Syntactic ambiguity and lexical ambiguity. In syntactic ambiguity, an expression can be decomposed into its constituents in more than one way. In lexical ambiguity, words have more than one meaning. Lexical ambiguity is one of the features that distinguish natural languages from formal languages. There is no lexical ambiguity in a programming language. Is lexical ambiguity an obstacle to modeling natural languages? In other words, does it have adverse effects on artificial intelligence studies? In this work, with reference to contemporary literature, it is stated that lexical ambiguity, especially polysemy, is not a defect but an advantage for cognitive systems.

Keywords: ambiguity, polysemy, natural language processing, cognitive economy, the small-world hypothesis.

Öz: Doğal dilleri anlayabilen bir sistemin temel özellikleri neler olmalıdır? Hangi engellerin aşılması gerekmektedir? Doğal dil anlama süreçleriyle ilgili önemli sorunlardan birisi de muğlaklık olarak görülmektedir. Muğlaklığın iki türünden bahsedilebilir: Sözdizimsel muğlaklık ve sözcüksel muğlaklık. Sözdizimsel muğlaklık durumunda verilen bir ifade bağlama göre birden fazla şekilde sözdizimsel parçalarına ayrılabilir. Muğlaklığın diğer bir türü ise sözcüksel muğlaklıktır. Bu durumda, kelimeler birden fazla anlama sahiptirler. Sözcüksel muğlaklık doğal dilleri biçimsel dillerden ayıran özelliklerden birisidir. Bir programlama dilinin söz

† Bu makale, birtakım ekleme ve düzenlemelerle birlikte, 2022 yılında, Orta Doğu Teknik Üniversitesi, Sosyal Bilimler Enstitüsü, Felsefe Anabilim Dalına sunduğum *A Defense of Meaning Eliminativism: A Connectionist Approach* adlı, yayımlanmamış doktora tezinin bir bölümünden uyarlanmıştır.

Author Info: Tolgahan TOY

Middle East Technical University, Faculty of Arts and Science, Department of Philosophy, 06800 Ankara, TÜRKİYE. E-mail: tolgahantoyboun@gmail.com

To Cite This Paper: Toy, T. (2023). "Sözcüksel Muğlaklığın Doğal Dil İşleme Sistemlerindeki Yeri." *MetaZihin*, 6(1): 17-31.

dağarcığındaki ifadeler için muğlaklıktan bahsetmek mümkün değildir. Sözcüksel muğlaklık doğal dillerin modellenmesinin önünde bir engel midir? Diğer bir deyişle doğal dillerdeki anlamsal muğlaklığın yapay zekâ çalışmalarını kötü yönde etkilediği söylenebilir mi? Bu çalışmada güncel literatür referans alınarak sözcüksel muğlaklığın özellikle de çokanlamlılığın doğal diller için bilişsel açıdan aslında bir avantaj olduğu ifade edilmektedir.

Anahtar Kelimeler: muğlaklık, çokanlamlılık, doğal dil işleme, bilişsel ekonomi, küçük dünya hipotezi

1. Giriş

Milyonlarca kitabı, makaleyi okuyup anlayabilecek, özetleyebilecek, sorulan sorulara cevap verebilecek bir sistem yapay zekâ çalışmalarının zirvesi olarak görülebilir. İnsan dilini kusursuz anlayabilen¹ mekanik bir sistem geliştirilebilir mi? Rene Descartes'in bu soruya yanıtı olumsuzdur. Descartes mekanik bir sistemin sistematik bir şekilde yeni cümleleri anlamlandıramayacağını, ifade edemeyeceğini iddia etmektedir.

(...) düşüncelerimizi başkalarına açıklamak için bizim yaptığımız gibi kelimeleri veya diğer işaretleri birleştirerek kullanamazlar. Çünkü bir makinenin kelimeleri aktaracak şekilde yapılmış olabileceği ve hatta organlarında bir değişikliğe neden olan bedensel eylemlerle ilişkili olarak, mesela ona dokunulduğunda, bazı sözler söyleyeceği pekâlâ tasavvur edilebilir. Belirli bir yerde, kendisine ne söylenmek istendiğini sorabilir; başka bir yerine dokunulursa inciniyor diye haykırabilir. Ama en akılsız insanların bile yapabildiği gibi karşısında söylenen her şeyin anlamına yanıt vermek için kelimeleri ve işaretleri çeşitli biçimlerde düzenleyemez. (Descartes, 1637/1968: 73-74)

Diğer taraftan, Alan Turing ise doğal dili anlayabilen mekanik bir sistemin mümkün olduğunu iddia etmektedir (Turing, 1950: 442). Daha da ileri giderek dili deneyimleyerek öğrenebilen bir sistemden bahsetmektedir:

Yapılacak en iyi şey makineye paranın satın alabileceği en iyi organları sağlamak ve ardından ona İngilizce anlamayı ve konuşmayı öğretmek olurdu. (Turing, 1950: 460)

Yapay zekâ çalışmalarındaki yükselişle birlikte artık soru doğal dilleri anlayabilen mekanik bir sistemin mümkün olup olmadığının ötesine geçmiştir. Yeni sorun böyle

¹ *Anlamak* terimi John Searle'ün yaptığı güçlü/zayıf yapay zekâ ayrımını çağrıştırmaktadır. Searle'e göre insanların dilsel davranışlarını taklit edebilen salt sözdizimsel bir sistem herhangi bir doğal dili anlayabilecek güçte değildir. Daha doğrusu anlamın ortaya çıkabilmesi için bir dizi biçimsel işlemden fazlası gerekmektedir (Searle, 1980). Ancak bizim çalışmamız dil anlama sürecinin taklit edilmesi ve *gerçekten* dil anlama arasında yapılan ayırım konusunda tarafsız bir tutum sergilemektedir. Doğal dil anlama teriminin kullanılmasının nedeni, doğal dil anlama (NLU) alanının doğal dil işleme (NLP) alanının yapay zekâ-tam (AI-complete) olarak görülen bir alt dalı olmasıdır (Shapiro, 1992, s. 56). Bu anlamıyla doğal dil anlama çalışmaları Searle'ün bahsettiği güçlü/zayıf yapay zekâ ayrımıyla ilişkili değildir.

bir sistemin nasıl mümkün olacağıdır? Doğal dilleri anlayabilen bir sistemin temel özellikleri neler olmalıdır? Hangi engellerin aşılması gerekmektedir?

Bu çalışmada bizi ilgilendiren ise doğal dil anlama sistemleri için bir sorun olarak görülen sözcüksel muğlaklığın bilişsel durumudur. Doğal dil anlama süreçleriyle ilgili önemli sorunlardan birisi olarak görülen muğlaklığın iki türünden bahsedilebilir: Sözdizimsel muğlaklık ve sözcüksel muğlaklık. Sözdizimsel muğlaklık durumunda verilen bir ifade bağlama göre birden fazla şekilde sözdizimsel parçalarına ayrılabilir. Mesela, “sarışın çocuk ve kadın köye gitti” ifadesi sözdizimsel açıdan iki farklı şekilde incelenebilir:

[[[sarışın çocuk] [[ve] [kadın]]] [[köye] [gitti]]]

[[[sarışın] [[çocuk] [[ve] [kadın]]]] [[köye] [gitti]]]

İlk incelemede çocuğun sarışın olduğu belirtilirken, kadının rengi ile ilgili herhangi bir bilgi verilmemektedir. İkinci incelemede ise çocuğun da kadının da sarışın oldukları ifade edilmektedir. Dolayısıyla, bu ifadeyle karşılaşan bir sistemin bağlama göre bu iki incelemeden birisini seçmesi gerekmektedir. İfadelerin uzunluğu, farklı sözdizimsel ilişkiler göz önüne alındığında, doğru sözdizimsel incelemeyi saptamak zorlaşmaktadır.

Sözcüksel muğlaklık durumunda ise kelimeler birden fazla anlama sahiptirler. Aşağıdaki örneklerde “çay” kelimesinin farklı anlamları görülmektedir:

- Çay toplamaya gittik.
- Çay içiyoruz.
- Çay’a düştü.
- Çay’a zam geldi.

Bir açıdan bakıldığında, sözcüksel muğlaklık bu ifadelerin anlamlarının hesaplanmasını zorlaştırmaktadır. Bu yaklaşıma göre formel dillerde olduğu gibi muğlaklıktan arınmış ifadeler mekanik bir sistem için daha uygun görünmektedir. Bu makalede ise sözcüksel muğlaklığın sorun oluşturmayıp aksine faydalı olduğu iddia edilmektedir. Sırasıyla, bir sonraki bölümde sözcüksel muğlaklığın neden bir sorun olarak görüldüğü kısaca özetlenmektedir. Çalışmanın ana bileşenleri olan üçüncü ve dördüncü bölümlerde muğlaklığın bilişsel bir sistem için faydaları aktarılmaktadır. Beşinci bölümde ise güncel dil anlama sistemleri referans alınarak, sözcüksel muğlaklığın neden bir sorun oluşturmadığı açıklanmaktadır.

2. Bir Sorun Olarak Sözcüksel Muğlaklık

Bir taraftan bakıldığında, kelimelerin birden fazla anlamının olması doğal dillerin bir kusuru olarak görülebilir (Wasow, Perfors ve Beaver, 2005: 269-270). Doğal diller insan evriminin bir ürünüdür. İnsan evriminin bir amaç doğrultusunda hareket etmediği ve tesadüfi mutasyonlarla şekillendiği düşünüldüğünde, bilişsel yeteneklerimizin mükemmel olmaması gayet doğal görünmektedir. Tıpkı gözlerimizin ışığa, kulaklarımızın sese duyarlılık konusunda mükemmel olmamaları gibi dil anlama yeteneğimizin de iletişim açısından mükemmel olmaması gayet olağandır. Noam Chomsky muğlaklığı işaret ederek, işlevselci yaklaşımların iddia ettiğinin aksine dilin iletişim için tasarlanmadığını düşünmektedir. Chomsky'e göre insan dilinin asıl görevi bilişsel yapıımızdaki diğer birimlerle uyumlu olmasıdır (Chomsky, 2002, s. 106-108).

Bir benzetme yapmaya çalışalım: Vücudun başka bir organını, örneğin karaciğeri ele alalım. Karaciğerin İtalya'da yaşamak için kötü tasarlandığını keşfedebilirsiniz çünkü İtalya'da insanlar çok fazla şarap içmekte ve her türlü karaciğer hastalığına yakalanmaktadır. Bu nedenle, karaciğer işlevsellik açısından iyi tasarlanmamıştır. Diğer taraftan, karaciğer dolaşım sistemi, böbrek vb. organlarla etkileşim için güzel bir şekilde tasarlanmış olabilir... Dilin iletişim için kullanılması bir tür epifenomen olabilir. Yani sistem nasıl gelişti, gerçekten bilmiyoruz. Ve sonra şunu sorabiliriz: İnsanlar bunu nasıl kullanıyor... Birbirimizi asla yanlış anlamadığımızdan emin olmak istiyorsanız, dil bu amaç için iyi tasarlanmamıştır çünkü belirsizlik gibi özelliklere sahiptir. Bu nedenle, sistem birçok işlevsel açıdan iyi tasarlanmamıştır. Ancak tamamen ayrı bir soru vardır: Etkileşimde bulunması gereken iç sistemler açısından iyi tasarlanmış mıdır? Bu farklı bir bakış açısı ve yeni bir soru: Minimalist Programın cevaplamaya çalıştığı soru da budur. (Chomsky, 2002: 107-108)

Sözlüksel muğlaklık yapay zekâ çalışmalarında başından itibaren sorun oluşturmuştur. Yehoshua Bar-Hillel bu soruna işaret ederek, tıpkı Descartes gibi, doğal dil işleyebilen mekanik bir sistemin mümkün olmadığını iddia etmektedir. İçinde muğlak sözcükler barındıran bir ifadenin anlamının hayal edilebilir hiçbir bilgisayar tarafından hesaplanamayacağını iddia etmektedir. Ona göre, "çay geldi" ifadesinde "çay" ile ne kastedildiği bir sözlük ve belirli kurallar ile belirlenemez (Bar-Hillel, 1964).

Hesaplama karmaşıklığı açısından sorunu şöyle formüle edebiliriz. n -kelimelik bir cümle düşünelim. Bu cümledeki her bir kelimenin ortalama m farklı anlamı olsun. Bu durumda, cümlelerin n^m farklı olası anlamı bulunmaktadır. n ve m sayıları arttıkça hesaplama işlemi daha da karmaşıklaşmaktadır (Falkum & Vicente, 2015: 3). Daha da kötüsü, dil kullanıcıları bağlam içinde kelimelere yeni anlamlar kazandırmaktadırlar (Pustejovsky, 1996: 42-46). Mesela, Heideggerci bir filozof "varlık kendini açtı" derken, "açmak" kelimesi sözlükte bulunmayan yeni bir anlama sahip olmaktadır. Eğer kelimelerin alabileceği anlamlar sayılabilir değilse bu durumda olası anlamlar arasından doğru anlamı seçebilecek bir hesaplama pratik görünmemektedir (Clark, 1983: 301; Cohen, 1985: 132,134).

3. Sözcüksel Muğlaklığın Bilişsel Ekonomiye Katkıları

Yukarıda bazı örnekleri sunulan yaklaşıma göre sözcüksel muğlaklık, anlamın bileşimsel bir şekilde hesaplanmasını oldukça zor bir hale getirmektedir. Diğer taraftan, muğlaklığın giderilmesi insanlar için kolaydır. Günlük iletişimde sürekli üstesinden geldiğimiz bir durumdur (Taylor, 2003: 647; Wasow, Perfors ve Beaver, 2005: 272; Wasow, 2015: 30, 32, 37-42). Mesela, birisine “çaydan geliyoruz” dediğinizde, ifadeyi hiçbir zorluk çekmeden, kullanıldığı bağlama göre anlamaktadır. Sözcüksel muğlaklığın bir istisna değil, doğal diller için bir norm olduğu (Cohen, 1985: 131-132; Clark, 1983: 299-300, 305) göz önünde bulundurulduğunda, yukarıdaki yaklaşımın gözden geçirilmesi gerekmektedir.

Literatürde sözcüksel muğlaklığı sorun olarak görmeyen bir yaklaşım da mevcuttur. Bu yaklaşıma göre sözcüksel muğlaklık bir sorun olmaktan ziyade, iletişim açısından istenen bir özelliktir (Zipf, 1949: 27; Cohen, 1985: 135; Harris, 1994: 222-223). Bu yaklaşımın kökenleri George Kingsley Zipf’in çalışmalarına dayanmaktadır. Zipf, sözcüksel muğlaklığı, insan davranışlarının arkasındaki mekanizmalardan biri olarak gördüğü bilişsel ekonomi üzerinden açıklamaktadır (Zipf, 1949: 19). Zipf için hem bireysel hem de grup davranışlarında birbirleriyle çatışan iki ekonomi bulunmaktadır. Çatışmayı çözen denge, grubun veya bireyin davranışını belirlemektedir. Dilsel iletişimde, konuşanların ve dinleyicilerin çıkarları çatışmaktadır. Konuşanlar birden fazla anlamı tek bir sözcükle ifade etmek isterken, dinleyenler belirsizliği minimum düzeye düşürmek, dolayısıyla daha az çaba harcamak isterler (Zipf, 1949: 20-22).

Konuşmacı açısından bu durumu şöyle düşünebiliriz. Yorgun birisinin sürekli “şey” kelimesini kullandığını düşünelim. Zipf’in bilişsel ekonomi kuramına göre konuşmacı “şey” kelimesini kullanarak, enerji sarfiyatını en aza indirmektedir. Tek heceli ve kolay telaffuz edilebilen “şey” kelimesinin çok sık kullanılmasının ve çok fazla anlama gelmesinin nedeni Zipf’e göre bilişsel ekonomidir (Zipf, 1949: 30-31). Diğer taraftan dinleyici ise en az çaba ile en fazla şeyi anlamak istemektedir. Dolayısıyla, ifadelerin net ve anlaşılır olmasını arzulamaktadır. Mesela, devamlı “şey” diyen bir konuşmacıyı dinlemek istemez.

Zipf’in kuramında çatışmaların çözümü, bireylerin davranışlarını belirlemektedir. Konuşmacılar ve dinleyiciler arasındaki çatışma iki tarafın da ihtiyaçlarını karşılayan makul miktarda bir belirsizlikle çözülmektedir (Zipf, 1949: 27-28).

Çağdaş literatürde Zipf’in bilişsel ekonomi kuramına enformasyon kuramı üzerinden yaklaşılmaktadır. Bağlam yeteri kadar bilgilendirici olduğunda, optimum muğlaklığın iletişim için faydalı olduğu kanısına varılmaktadır (Piantadosi, Tily ve Gibson, 2012:

282). Dolayısıyla, konuyu daha iyi incelemek için enformasyon kuramına başvurabiliriz. Claude Shannon'un yaklaşımında enformasyon, vericiden alıcıya giderken oluşan belirsizlik üzerinden ölçülmektedir (Shannon, 1948: 394). Shannon, gerçek mesajın olası mesajlar arasından seçildiğini, dolayısıyla sistemin bu olasılıklarla başa çıkabilecek yapıda olması gerektiğini aktarmaktadır. Çünkü, sistem hangi konuşmacının hangi mesajı vermek istediğini kendisine gönderilen veri üzerinden çözmektedir. Dolayısıyla, alıcının sistemi, verinin karşılık gelebileceği olasılıklar üzerinden işlem yapmaktadır (Shannon, 1948: 379).

Dilsel iletişim açısından bakıldığında bir kelimenin muğlaklığını Shannon'un geliştirdiği aşağıdaki entropi formülüyle hesaplayabiliriz.

$$H[\text{Anlam}] = - \sum_{\text{anlam} \in \text{Anlam}} P(\text{anlam}) \log P(\text{anlam})$$

Ancak, ihtiyacımız olan şey koşullu entropidir. Çünkü dinleyiciler kelimeleri belirli bir bağlam içinde işlemektedirler.

$$H[\text{Anlam} | \text{Bağlam}] = - \sum_{\text{bağlam} \in \text{Bağlam}} P(c) \sum_{\text{anlam} \in \text{Anlam}} P(\text{anlam} | \text{bağlam}) \log P(\text{anlam} | \text{bağlam})$$

Koşullu entropi, koşulsuz entropiden daha büyük olamayacağı için anlamın bağlama göre değişmediği bir dil ekonomik değildir (Piantadosi, Tily ve Gibson, 2012: 283). Bu durumu şöyle açıklayabiliriz: Dil kullanıcıları her zaman bir bağlam içerisindedirler. Sadece çevremizde bulunan somut nesnelere konuşurken değil, soyut meseleler üzerine tartışırken de dili bir bağlam içerisinde kullanırız (Falkum ve Vicente, 2015: 90; Rayo, 2013: 655-656). İfadelerin anlamlarının bağlama göre değişmediği bir dil düşünelim. Bu durumda $p(\text{anlam} | \text{bağlam}) = p(\text{anlam})$ olacaktır. Bu da koşullu ve koşulsuz entropinin aynı olacağı anlamına gelmektedir. Eğer kelimelerin anlamı bağlama göre değişirse koşullu entropi koşulsuz entropiden daha küçük olacaktır. Yani, aynı mesajı göndermek için daha az birim (bit) gerekecektir.

Daha spesifik olarak konuşmacının ve dinleyicinin harcadıkları çaba şu şekilde hesaplanabilir. Konuşmacı uygun kelimeleri ararken bir çaba harcamaktadır:

$$H[\text{Kelime}] = - \sum_{\text{kelime} \in \text{Kelime}} p(\text{kelime}) \log p(\text{kelime})$$

Dinleyici ise verilen kelimeler için doğru anlamı bulmaya çalışırken çaba harcamaktadır:

$$H[\text{Anlam} | \text{Kelime}] = - \sum_{\text{kelime} \in \text{Kelime}} p(\text{kelime}) \sum_{\text{kelime} \in \text{Kelime}} p(\text{anlam} | \text{kelime}) \log p(\text{anlam} | \text{kelime})$$

Konuşmada harcanan toplam enerji ise şöyle hesaplanmaktadır (Ferrer i Cancho ve Solé, 2003; Solé ve Seoane, 2015).

$$\Omega(\lambda) = \lambda H[\text{Anlam} | \text{Kelime}] + (1 - \lambda) H[\text{Kelime}]$$

Dilde belirsizlik yoksa, $H[\text{Anlam} | \text{Kelime}]$ 0'dır. Bu durum, dinleyicilerin ifadelerdeki muğlaklığı gidermek için herhangi bir çaba harcamalarına gerek olmadığı anlamına gelmektedir. Ancak böyle bir durum konuşmacılar için zor olacaktır. Öte yandan, eğer konuşmacılar sadece bir kelime kullanarak istediği herhangi bir şeyi ifade edebiliyorlarsa, $H(\text{Kelime})$ 0 olur (Solé & Seoane, 2015: 21-22). Dolayısıyla, konuşmacılar hiçbir çaba sarfetmeden istediği her şeyi ifade edebilmektedirler. Özetle, Ω formülündeki λ 'ya bağlı olarak, konuşmacının harcadığı enerji, $H(\text{Kelime})$ ya da dinleyicinin harcadığı enerji, $H(\text{Anlam} | \text{Kelime})$ daha yüksektir. Ferrer i Cancho ve Solé, toplam harcanan enerjiyi, yani Ω 'yı en aza indirmeyi amaçlayan bir algoritma geliştirmişlerdir. Bu algoritmaya göre, $\lambda=0.5$ olduğunda, Ω minimize edilmektedir. Bu da Zipf'in, konuşanın ve dinleyenin çabasını eşit olarak gören yaklaşımına katkı sağlamaktadır (Ferrer i Cancho ve Solé, 2003: 789; Solé ve Seoane, 2015: 25).

$\lambda < 1 - \lambda$ olduğunda iletişim yok olmaya doğru gitmektedir. Diğer taraftan, $\lambda > 0.5$ eşitsizliği ise yapay dillerdekine benzer bir durum yaratmaktadır (Ferrer i Cancho ve Solé, 2003: 789-790). İnsan dillerinde ise makul düzeyde bir belirsizlik mevcuttur (Solé ve Seoane, 2015: 25; Ferrer i Cancho ve Solé, 2003: 790-791).

Ayrıca İngilizce, Felemenkçe ve Almanca üzerine yapılan çalışmalarda bir kelimeye karşılık gelen anlamların sayısı ile o kelimeyi telaffuz etmenin kolaylığı arasında ilişki bulunduğu görülmektedir (Piantadosi, Tily ve Gibson, 2012: 285-287). Catherine L. Caldwell-Harris, kolay telaffuz edilebilen, sık kullanılan kelimelerin teknolojiye ve kültürel değişime bağlı olarak ortaya çıkan yeni alanlarda daha çok kullanıldığını aktarmaktadır (Harris, 1994, p. 223).

Özetle, doğal dillerdeki muğlaklık, dil kullanıcılarının çevre kaynaklarını daha iyi kullanma becerileri ile uyumludur. Bağlam yeteri kadar enformasyon sağladığında, dil kullanıcılarının, harcadıkları çabayı en aza indirmeleri adaptasyonel bir davranıştır. Kelimelerle ilişkilendirilen anlamların sayısı ve kelimelerin kolay telaffuz edilebilirliği arasındaki ilişki enerji sarfiyatını etkileyerek, sistemin hayatta kalmasını sağlamaktadır.

Beşinci bölümde açıklandığı gibi güncel dil anlama sistemleri henüz dil kullanımımızı diğer bilişsel özelliklerimizle ilişkilendirme konusunda yeteri kadar güçlü değildir. Ancak, bütünsel bir yapay zekâ sistemine doğru ilerledikçe sistemin çevreyle etkileşimi, çevreye adaptasyonu daha fazla önem kazanacaktır. Bu açıdan sistemin

çevrede halihazırda bulunan kaynakları, yani bağlamı, daha ekonomik kullanması sistem için bir avantajdır. Sonuç olarak, dilsel ifadelerin bağlama göre farklı anlamlar aldığı bir sistem yapay zekâ çalışmalarına katkı sağlamaktadır

4. Çokanlamlılık ve Küçük Dünya Hipotezi

Nelson Goodman'ın da belirttiği gibi, sözdizimsel yapı anlamsal yapı kadar yoğun olmadığına muğlaklık ortaya çıkmaktadır (Goodman, 1968: 12). Daha doğrusu, dil gibi görece seyrek, kesikli bir yapı ile zengin anlam ağı arasındaki ilişki muğlaklık üzerinden kurulmaktadır. Sözcüksel muğlaklığın özel bir türü olan çokanlamlılık bu açıdan oldukça önemlidir. Doğal dillerdeki çokanlamlılık, anlam ağının farklı noktalarını birbirlerine bağlamaktadır. Bu noktada, ağ kuramındaki küçük dünya hipotezi önemli bir rol oynamaktadır.

Eğer bir kelimenin anlamları arasında bağlantı bulunmuyorsa, kelime eşadlılık durumunu örneklendirmektedir. Kelimenin anlamları arasında ilişki bulunuyorsa, bu durum çokanlamlılığı örneklendirmektedir. "Açmak" kelimesinin aşağıdaki kullanımları çokanlamlılığı örneklendirmektedir.

- Kapıyı açtım.
- Televizyonu açtım.
- Konu konuyu açtı.
- Ülkesini dünyaya açtı.

"Açmak" kelimesinin bu örneklerdeki anlamları birbirleriyle ilişkilidir. Farklı alanlar "açmak" kelimesi üzerinden birbirlerine bağlanmaktadır. Çokanlamlılık anlam ağımıza "küçük dünya" yapısı vermektedir (Solé, Corominas-Murtra, Valverde ve Steels, 2010: 22; Solé ve Seoane, 2015: 16). Küçük dünya durumunda, çok sayıda düğümden oluşan bir ağda bir düğümden diğer bir düğüme az sayıda adımda ulaşabilmektedir (de Sola Pool ve Kochen, 1978; Milgram, 1967). Düğümler, birbirlerine bağlanarak ağ oluşturulan birimlerdir. Bunlar herhangi bir nesne olabilirler. Bir ağ insanlardan, nöronlardan, atomlardan, milletlerden vs. oluşabilmektedir. Ağa, küçük dünya yapısını veren düğümlerin birbirleriyle iyi bağlanmış olmalarıdır.

Örnek olarak iyi yapılandırılmış bir ağ olan Amerikan toplumunda herhangi bir insandan diğerine 2 kişi aracılığı ile ulaşabilmektedir. Diğer taraftan, birbirlerine yeteri kadar bağlı olmayan sadece 1000 kişilik bir toplulukta, bir insandan diğerine ulaşabilmek için en az 2 kişi bulunmaktadır (de Sola Pool ve Kochen, 1978: 42). Aynı şekilde, anlamlar birbirlerine iyi bir şekilde bağlandığında, bir anlamdan diğerine geçiş kolaylaşmaktadır. Mesela, "kitap" kelimesinin birbirleriyle ilişkili anlamlarını düşünelim. Fiziksel bir nesne olan kitap yapısında ağaç kullanılmaktadır. Diğer taraftan, soyut bir nesne olan kitap ise eğitim kavramıyla ilişkilidir. Ağaç ve eğitim gibi uzak görünen kavramların birbirleriyle olan ilişkisi "kitap" kelimesinin çokanlamlılığı

üzerinden görülmektedir. Dolayısıyla, çokanlamlılık anlam ağının küçük dünya özelliğine sahip olduğunun bir göstergesidir.

Anlamı listelemek yerine bir ağ ile temsil etmek anlamın en ince tonlarıyla, nüanslarla başa çıkmanın en iyi yoludur (Norvig ve Lakoff, 1987: 205; Brugman ve Lakoff, 1988: 477; Fellbaum, 2013: 3). Güçlü bir anlam ağı dilsel iletişim için oldukça önemlidir. Ayrıca, dil edinimi sürecinde, çocukların yeni anlamları bir ağ üzerinden öğrendikleri görülmektedir (Srinivasan ve Rabagliati, 2015: 146-147; Fellbaum, 2013: 5). Anlamın güçlü, sık dokunmuş yapısına karşılık, dil az sayıda, ayrı kelimelerden oluşmaktadır. Çokanlamlılık, bu iki yapı arasında köprü vazifesi görmektedir.

Sonuç olarak, insanların zihinsel dünyalarının zenginliğine, karmaşıklığına karşılık dilsel ifadeler ayrı ve basittir. İnsan davranışlarını sistematik bir şekilde taklit etmeyi amaçlayan bir sistemin, az sayıda sembol ile zengin zihin dünyamızı eşleyebilmesi gerekmektedir. Ayrı dilsel ifadeler üzerinden yoğun anlam ağları oluşturmak, sistemin bilimsel bir makaleyi okuyup duysal veriler ile eşleyebilmesi, farklı konuları birbirlerine bağlayabilmesi gibi özellikler kazanmasına katkı sağlayacaktır. Çok anlamlılık, örnek olarak "kütle" ya da "güç" kelimesinin önceki anlamlarına çok yakın olan yeni anlamlar kazanması için bir gerekliliktir.

Buradaki gereklilik iki açıdan önemlidir. Birincisi herhangi iki anlam arasında yeni bir anlamın ortaya çıkmasının önünde hiçbir engel yoktur. Dolayısıyla, anlamların gerçel sayılara karşılık gelmediğini düşünsek bile sıkı yapıları dolayısıyla en azından rasyonel sayılara karşılık geldiğini söyleyebiliriz. Rasyonel sayılara karşılık gelebilecek ifadelerden oluşan bir dil gerçekçi görünmemektedir. Bunun yerine halihazırda bulunan kelimelerin esnek bir şekilde bir anlamdan yakınında bulunan başka bir anlama geçebilmesi oldukça önemlidir. İkincisi, anlamsal yapı sadece anlamlardan değil aynı zamanda anlamlar arasındaki ilişkilerden oluşmaktadır. Çok anlamlılık olmadığında bu ilişkiler yeterli bir şekilde temsil edilmemiş olacaktır.

5. Muğlaklık ve Yapay Sınır Ağları

Önceki iki bölümde sözcüksel muğlaklığın yapay zekâ sistemlerine ne tür katkılar sağlayabileceği aktarıldı. Sözcüksel muğlaklık bilişsel bir sistemin kaynaklarını daha ekonomik kullanmasına ve zengin bir anlam dünyasıyla başa çıkabilmesine olanak vermektedir. Peki bu faydalarına karşılık muğlaklığın doğal dil anlama sistemleri için oluşturduğu sorunlardan bahsedebilir miyiz? Bu bölümde muğlaklığın yapay zekâ sistemleri için herhangi bir sorun oluşturmadığı belirtilmektedir.

İkinci bölümde sözcüksel muğlaklığı sorun olarak gören yaklaşımlardan bahsedilmişti. Bu görüşlerin ortak yanı, ifadelerin anlamlarının bileşimsel olarak hesaplanmasıdır. Diğer bir deyişle, bu görüşlere göre bir ifadenin anlamı parçalarının anlamlarının bir fonksiyonudur (Frege, 1879/1967: §. 9; Hodges, 2001: 7; Montague, 1974: 227; Pagin ve Westerståhl, 2010: 250; Partee, 1984: 281-282; Szabó, 2012: 75; Szabó, 2017). Mesela,

“yolun sonuna geldik” ifadesinin anlamını hesaplamak için “yol”, “son” “gelmek” kelimelerinin anlamları ve cümlenin sözdizimsel yapısı gerekmektedir. Ek olarak, bu kelimelerin bağlam içerisinde hangi anlamlarının kullanılacağını belirleyen bir yonteme ihtiyacımız vardır. Dilsel ifadelerin uzunluğu ve dilin birbirinden çok farklı bağlamlarda kullanılabilceği göz önüne alındığında, muğlaklığın bileşimsel yaklaşım için sorun oluşturduğu görülmektedir. Ancak, güncel doğal dil işleme sistemlerinde bileşimsel bir yaklaşım yerine bağlantıcı algoritmalar kullanılmaktadır.

Anlamı bir ağ olarak gören bağlantıcı sistemler muğlaklığın bağlam içerisinde giderilmesi konusunda oldukça başarılıdır. Özyinelemeli derin sinir ağları ya da dönüştürücü dil modelleri kullanılarak, kelimelerin bağlam içerisindeki anlamları hesaplanmaktadır (Peters vd., 2017; Peters vd., 2018). Oldukça başarılı olan bu yaklaşımlar, kelimelerin anlamlarını, kelimenin solunda ve sağında bulunan kelimeleri hesaba katarak belirlemektedirler. Bu yöntem ile başarılı bir şekilde, az sayıda kelime ile çok boyutlu, sonsuz bir anlam uzayının farklı bölgelerine gönderimler yapılabilmektedir.

Güncel bağlantıcı dil işleme sistemlerini, (Vaswani, vd. 2017)’de tanıtılan *dikkat* temelli dönüştürücü dil modelleri üzerinden özetleyebiliriz. Model, girdi olarak bir ifade (kelimelerden, harflerden oluşan bir sekans) verildiğinde, çıktı olarak yine bir ifade vermektedir. Model, çıktı girdinin başka bir dildeki karşılığı olduğunda çevirmen, çıktı girdinin biçimsel bir dildeki karşılığı olduğunda anlamsal çözümleyici, çıktı girdide sorulan sorunun cevabı olduğunda diyalog sistemi olarak çalışmaktadır. Model iki temel birimden oluşmaktadır: Kodlayıcı ve kod çözücü (Sutskever, Vinyals ve Le, 2014).

Harf ya da kelime sekansı olarak verilen girdi kodlayıcıya gönderilmektedir. Bu işlem iki aşamada yapılmaktadır. Sekansın kelimelerden oluştuğunu düşünürsek, sözlüğümüzde her bir kelime için d boyutlu bir vektör bulunmaktadır. Dilsel ifadelerde kelime sırasının değişmesi anlamı değiştirdiği için sinusoid fonksiyonu kullanılarak kelimelerin sekanstaki yerini temsil eden vektörler elde edilmektedir. $i < d/2$ olmak şartıyla ve n bir hiperparametre olarak alındığında her bir vektörün $2i$ ve $2i+1$ boyutlarındaki değerleri aşağıdaki şekilde hesaplanmaktadır:

$$PE(konum, 2i) = \sin\left(\frac{konum}{n^{2i/d}}\right)$$

$$PE(konum, 2i + 1) = \cos\left(\frac{konum}{n^{2i/d}}\right)$$

Kelime vektörleri, konumsal bilgi veren vektörlerle birleştirilerek kod çözücüye gönderilmektedir. Kod çözücü sekansın hangi noktalarına dikkat etmesi gerektiğini hesaplamak için bu vektörleri *sorgu*, *anahtar* ve *değer* vektörlerine dönüştürmektedir. Sorgu ve anahtar vektörleri öğrenilmiş bir uzaya transfer edildikten sonra birbirleriyle çarpılmaktadır. Bu yöntemle, sekansın içindeki her bir kelimenin diğer kelimelerle ne kadar yakın olduğu belirlenmektedir. Daha sonra bu değerler, softmax fonksiyonu kullanılarak, 0 ve 1 arasındaki ağırlıklara dönüştürülmektedir. (Ek olarak, öğrenme sırasında türev alımını kolaylaştırmak için her bir değer \sqrt{d} değerine bölünmektedir.) Elde edilen ağırlıklar, değer vektörleriyle çarpılarak dikkatin, sekansın hangi bölümlerine verilmesi gerektiğini belirten bir matris elde edilmektedir. Sorgu, anahtar ve değer vektörlerinden oluşan matrisleri sırasıyla Q , K ve V harfleriyle temsil edersek, yukarıda işlemleri aşağıdaki formülle özetleyebiliriz.

$$\forall q \in Q \text{ Dikkat}(q, K, V) = \left(\frac{q K^T}{e^{\frac{q K^T}{\sqrt{d}}}} \right) V$$

Kelimelerin sadece ikili ilişkileri değil, çoklu ilişkilerinin de hesaba katılması için bu işlem N defa tekrarlanmaktadır. Ayrıca sorgu, anahtar ve değer vektörleri daha küçük boyutlu vektörlere parçalanarak, yukarıdaki işlem bu vektörlerde paralel olarak uygulanmaktadır. Bu yolla, sekansdaki kelimelerin birbirleriyle ilişkileri farklı açılardan ele alınmış olmaktadır. Örnek olarak, kelimelerin cinsiyet ve coğrafya üzerinden ilişkileri paralel olarak hesaplanmaktadır. Son aşamada elde edilen vektörler tekrar birleştirilerek öğrenilmiş bir uzaya transfer edilmektedir.

Elde edilen vektörler, cümle başlangıcı anlamına gelen <SOS> sembolüne karşılık gelen vektör ile birlikte kod çözücüye gönderilmektedir. Kod çözücü kodlayıcı ile benzer bir yapıya sahiptir. Kod çözücünün verdiği değere softmax fonksiyona uygulanarak, hedef ifadenin ilk kelimesi üretilmektedir. Sistemin test edilmesi sırasında, elde edilen kelimeler, cümle sonu anlamına gelen <EOS> sembolüne ulaşana kadar kodlayıcıdan gelen değerle birlikte kod çözücüye gönderilerek işlem tekrarlanır.

Sorulan sorulara cevap verme, çeviri, anlamsal çözümleme gibi görevleri yapmakta oldukça başarılı olan bu sistem bileşimsel yaklaşımda olduğu gibi muğlaklığı bir sorun olarak görmemektedir. Kelimelere atadığı vektörler sözlükten seçtiğimiz anlamlar değildir. Dolayısıyla, verilen bir ifadenin anlamı parçalarının anlamı kullanılarak hesaplanmamaktadır. Anlam kelime vektörlerinin birbirleri ile çarpımları ve öğrenilmiş uzaylara transferleri sonucu ortaya çıkmaktadır. Sayıca az ve ayrık olan kelimeler, sekans içinde verildiklerinde, sistemi d boyutlu zengin bir uzayın farklı noktalarına götürebilmektedirler.

6. Sonuç

Sonuç olarak, doğal dillerdeki muğlaklık yapay zekâ çalışmalarının önünde bir engel değildir. Aksine, bilişsel bir sistem için katkı sağlamaktadır. Beşinci bölümde de ifade edildiği gibi güncel dil işleme modelleri muğlaklık konusunda oldukça başarılıdır. Ancak bu başarılar şu an için kelime ve söz öbeği düzeyindedirler. Ayrıca, sadece dilsel bağlam hesaba katılmaktadır. Dil kullanıcılarının maruz kaldığı diğer girdilerin, dil anlama edimine etkisi göz önünde bulundurulmamaktadır. Bu çalışmalar ifade düzeyinde, daha bağlamsal bir dil anlama modeli için temel oluşturmaktadırlar. Buradaki amaç, seyrek ve ayrık yapıda olan dil ile yoğun ve karmaşık yapıda olan anlam ağı arasında bağlam üzerinden bir ilişki kurabilmektir. Bunun başarılması yapay bir sistemin, dili insan düzeyinde kullanabilmesi için gerekli görünmektedir. Sözcüksel muğlaklığın, özellikle de çokanlamlılığın, bilişsel sistemlerimize katkısının daha iyi anlaşılması ve modellenmesi, doğal dil anlama çalışmalarına katkı sağlayacaktır.

Bu açıdan bakıldığında, üçüncü ve dördüncü bölümlerde tartışılan noktalar önemlidir. İlk olarak, az kelime ile daha çok şeyin ifade edilebilmesi/anlaşılabilmesi bilişsel süreçleri daha ekonomik hale getirmektedir. Kelime ya da genel olarak ifade sayısını arttırmak yerine, halihazırda bulunan kaynakları yani bağlamı kullanmak sistemin daha az enerji harcamasını sağlamaktadır. Dolayısıyla, bu özellik yapay ya da doğal bir sistemin çevreyle adaptasyonuna katkı sağlamaktadır. Ek olarak, muğlaklığın özel bir türü olan çokanlamlılık, sistemin güçlü bir anlam ağıyla başa çıkabilmesini sağlamaktadır. Farklı konular oldukça esnek bir şekilde semboller üzerinden birbirlerine bağlanmaktadır. Mesela, felsefe ve bilimler arasındaki ilişki "anne" kelimesi üzerinden anneler ve çocukları arasındaki ilişkiye benzetilebilmektedir. Dilin daha esnek kullanımları dünyayı yeni bakış açıları oluşturarak anlamamızı sağlamaktadır. Dolayısıyla, çokanlamlılık, yapay ya da doğal, dil anlayabilen herhangi bir sistemin dünya ile etkileşimini güçlendirmekte ve bütünselleştirmektedir.

7. Kaynakça

- Bar-Hillel, Y. (1964). "A demonstration of the nonfeasibility of fully automatic high quality machine translation." *Language and Information: Selected Essays on Their Theory and Application* içinde (s. 174-179). Reading, Mass.: Addison-Wesley.
- Baxter, D. L. (2016). "Hume on Space and Time." P. Russell (Der.), *The Oxford Handbook of Hume* içinde (s. 173-190). Oxford: Oxford University Press.
- Bonnen, C. A., ve Flage, D. E. (2000). "Descartes: The Matter of Time." *International Studies in Philosophy*, 32(4): 1-11.

- Brugman, C., ve Lakoff, G. (1988). "Cognitive Topology and Lexical Networks." G. W. Steven L. Small ve M. K. Tanenhaus (Der.), *Lexical Ambiguity Resolution: Perspectives from Psycholinguistics, Neuropsychology, and Artificial Intelligence* içinde (s. 477–508). San, Mateo, CA: Morgan Kaufmann.
- Chomsky, N. (2002). *On Nature and Language*. New York: Cambridge University Press.
- Clark, H. H. (1983). "Making Sense of Nonce Sense." G. B. D'Arcais, ve R. J. Jarvella (Der.), *The Process of Language Understanding* içinde (s. 297–331). John Wiley & Sons Ltd.
- Cohen, J. (1985). "A problem about Ambiguity in Truth-Conditional Semantics." *Analysis*, 45: 129–135.
- de Sola Pool, I., & Kochen, M. (1978). "Contacts and Influence." *Social Networks*, 1: 5–51.
- Descartes, R. (1637/1968). "Discourse on Method and the Meditations." Çev. F. E. Sutcliffe. London: Penguin Books Ltd.
- Descartes, R. (1985). *Meditations on First Philosophy*. J. Cottingham, R. Stoothoff, ve D. Murdoch (Düz.), *The Philosophical Writings of Descartes* (Cilt II) içinde, (s. 1-63). Cambridge: Cambridge University Press.
- Falkum, I. L., ve Vicente, A. (2015). "Polysemy: Current Perspectives and Approaches." *Lingua*, 157: 1-16.
- Fellbaum, C. (2013). "WordNet." C. A. Chappelle (Der.), *The Encyclopedia of Applied Linguistics* içinde (s. 707-728). Wiley/Blackwell.
- Ferrer i Cancho, R., ve Solé, R. V. (2003). "Least Effort and the Origins of Scaling in Human Language." *Proceedings of the National Academy of Sciences of the United States of America*, 100(3): 788–791.
- Garrett, D. (1981). "Hume's Self-Doubts about Personal Identity." *The Philosophical Review*, 90(3): 337-358.
- Garrett, D. (2008). "Hume's Theory of Ideas." E. S. Radcliffe (Der.), *A Companion to Hume* içinde (s. 41-57). Oxford: Blackwell Publishing .
- Giles, J. (1993). "The No-Self Theory: Hume, Buddhism, and Personal Identity." *Philosophy East and West*, 43(2): 175-200.

- Goodman, N. (1968). *Languages of Art: An Approach to a Theory of Symbols*. Hackett publishing.
- Harris, C. L. (1994). "Coarse Coding and the Lexicon." C. Fuchs ve B. Victorri (Der.), *Continuity in Linguistic Semantics* içinde (s. 205–229). Amsterdam: John Benjamins Publishing Company.
- Hume, D. (1960). *A Treatise of Human Nature*. Düz. L. A. Selby-Biggie. Oxford: The Calderon Press.
- Locke, J. (1999). *Essay Concerning Human Understanding*. Pennsylvania: Pennsylvania State University Press.
- Melamed, Y. Y. (2014). "What is Time?" A. Garrett (Der.), *The Routledge Companion to Eighteenth Century Philosophy* içinde (s. 232-244). London & New York: Routledge.
- Mermin, N. D. (2009). "What's Bad about This Habit." *Physics Today* 62: 8-9.
- Milgram, S. (1967). "The Small World Problem." *Psychology today*, 2: 60–67.
- Morrison, P. (1978). "Kant, Husserl and Heidegger on Time and Unity of 'Consciousness'." *Philosophical and Phenomenological Research*, 39(2): 182-198.
- Norvig, P. ve Lakoff, G. (1987). "Taking: A Study in Lexical Network Theory." *Annual Meeting of the Berkeley Linguistics Society*, 13: 195–206.
- Penelhum, T. M. (1976). "The Self in Hume's Philosophy." *The Southwestern Journal of Philosophy*, 7(2): 9-23.
- Peters, M. E., Ammar, W., Bhagavatula, C., ve Power, R. (2017). "Semi-Supervised Sequence Tagging with Bidirectional Language Models." *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, Volume 1: Long Papers* içinde (s. 1756–1765). Vancouver, Canada. Association for Computational Linguistics.
- Peters, M. E., Neumann, M., Iyyer, M., Gardner, M., Clark, C., Lee, K., ve Zettlemoyer, L. (2018). "Deep Contextualized Word Representations." *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1: Long Papers* içinde (s. 2227–2237). New Orleans, Louisiana. Association for Computational Linguistics.

- Piantadosi, S. T., Tily, H., ve Gibson, E. (2012). "The Communicative Function of Ambiguity in Language." *Cognition*, 122: 280–291.
- Pitson, A. E. (2002). *Hume's Philosophy of the Self*. London & New York: Routledge.
- Pustejovsky, J. (1996). *The Generative Lexicon*. MIT Press.
- Rayo, A. (2013). "A Plea for Semantic Localism." *Noûs*, 47: 647–679.
- Shannon, C. E. (1948). "A Mathematical Theory of Communication." *The Bell System Technical Journal*, 27: 379-423.
- Solé, R. V. ve Seoane, L. F. (2015). "Ambiguity in Language Networks." *The Linguistic Review*, 32: 5–35.
- Solé, R. V., Corominas-Murtra, B., Valverde, S. ve Steels, L. (2010). "Language Networks: Their Structure, Function, and Evolution." *Complexity*, 15: 20–26.
- Srinivasan, M. ve Rabagliati, H. (2015). "How Concepts and Conventions Structure the Lexicon: Cross-Linguistic Evidence from Polysemy." *Lingua*, 157: 124–152.
- Taylor, J. R. (2003). "Polysemy's Paradoxes." *Language Sciences*, 25: 637-655. [https://doi.org/10.1016/S0388-0001\(03\)00031-7](https://doi.org/10.1016/S0388-0001(03)00031-7)
- Turing, A. M. (1950). "Computing Machinery and Intelligence." *Mind*, 59: 433–60. DOI: 10.1093/mind/LIX.236.433
- Wasow, T. (2015). "Ambiguity Avoidance is Overrated." Susanne Winkler (Der.), *Ambiguity: Language and Communication* içinde (s. 29-48). De Gruyter.
- Wasow, T., Perfors, A. ve Beaver, D. (2005). "The Puzzle of Ambiguity." C. Orhan Orgun ve Peter Sells (Der.), *Morphology and the Web of Grammar: Essays in Memory of Steven G. Lapointe* içinde (s. 265–282). CSLI Publications.
- Waxman, W. (2008). "Hume and the Origin of Our Ideas of Space and Time." E. S. Radcliffe (Der.), *A Companion to Hume* içinde (s. 72-88). Oxford: Blackwell Publishing.
- Zipf, G. K. (1949). *Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology*. New York: Hafner Publishing Company.

==== This Page Intentionally Left Blank ====