



# JOURNAL OF SOCIAL AND HUMANITIES SCIENCES RESEARCH

Uluslararası Sosyal ve Beşeri Bilimler Araştırma Dergisi  
Open Access Refereed e-Journal & Refereed & Indexed

Article Type	Research Article	Accepted / Makale Kabul	26.01.2019
Received / Makale Geliş	17.11.2018	Published / Yayınlanma	27.01.2019

## MAKİNE ÇEVİRİSİNİN KISA TARİHÇESİ A BRIEF HISTORY OF MACHINE TRANSLATION

Dr. Öğr. Üyesi İnönü KORKMAZ

Trakya Üniversitesi, Edebiyat Fakültesi, Mütercim-Tercümanlık Bölümü, Edirne / TÜRKİYE,  
ORCID: 0000-0001-5601-1997

### ÖZET

Çevirinin farklı kültürlerden ve dil gruplarından bireyler arasındaki iletişimde önemli bir rolünün bulunması zaman içinde bu eylemin nasıl daha hızlı ve verimli bir biçimde gerçekleştirilebileceğine dair yolların aranmasına neden olmuştur. Günümüzden birkaç yüz yıl önce Descartes'ın tüm insanların kullanabileceği ortak bir dil sistemi kurma çabasının olası nedeni de bu olmuştur. Descartes'in ortak bir sistemde diğer dilleri birleştirme düşüncesi yirminci yüzyılda çoğunlukla askeri konularda istihbarat elde etmede kullanılmak üzere geliştirilen ve şifreleme ile şifreleri kırma işlevlerini yerine getiren makinelerin temel prensibini oluşturmaktaydı. Temelde belirli bir dilde kodlanarak yazılmış metinlerdeki bilginin kodlarının çözülmesi yoluyla bir başka dile kolayca aktarılması amaçlanıyordu. Bu düşünce Makine Çevirisinin (MÇ) temelini oluşturmuştur. Buradan hareketle de insan çevirmenlerin müdahalesinin en aza indirgenerek çeviri eyleminin bir nevi otomatikleştirilmesi hedeflenmişti. Ancak zaman içinde insan çevirmenlerin müdahalesinin nitelikli çevirilerin elde edilmesi için gerekliliğinin ortaya çıkmasıyla tamamen otomatik çeviri sistemleri arayışından insan çevirmenlerin çeviri sürecini kolaylaştırmanın ve hızlandırmanın yolları aranmaya başlamıştır. Böylece de makine çevirisi kavramı çeviri uygulamalarında Bilgisayar Destekli Çeviri olarak adlandırılan yeni bir kavramın ortaya çıkmasına hizmet etmiştir.

Bu çalışmanın amacı Makine Çevirisi kavramının tarihsel olarak belli başlı dönüm noktalarına işaret ederek gelişim sürecinden başlayarak günümüze kadar hangi evrelerden geçtiğini betimlemektir. Bu dönüm noktaları arasında ilk patenter ve ENIAC, Warren Weaver'in öngörüsü, ALPAC raporunun yayınlanması ve ilk ticari çeviri sistemlerinin ortaya çıkması makine çevirisinin tarihi için önemli olaylar olarak değerlendirilebilir. MÇ'nin geçirdiği evreler arasında ise makine çevirisinde kullanılan sistemlerden Kural Tabanlı, Örnek Tabanlı ve İstatiksel makine çeviri sistemleri bulunmaktadır. Ayrıca birden fazla sistemin bir arada kullanıldığı melez (hibrid) MÇ yaklaşımları da ortaya çıkmıştır. Günümüzde ise yapay zekâ sistemleri ile makine çevirisi sistemlerinin birleştirilmesine yönelik çalışmalar gündeme gelmiştir ve bu alana ilgi gün geçtikçe artmaktadır. Çalışmada ayrıca MÇ'nin gelişimi sürecinde karşılaşılan zorlukların yanı sıra gelişimini destekleyen faktörlere değinilmekte ve makine çevirisinin genellikle etkili olarak kullanıldığı alanlardan bahsedilmektedir. Son olarak ise çeviri uygulamalarının hangi yönde ilerlemekte olduğuna dair izlenimlere yer verilmektedir.

**Anahtar Kelimeler:** Makine Çevirisi, İstatiksel Makine Çevirisi, Kural-Tabanlı Makine Çevirisi, Örnek-Tabanlı Makine Çevirisi.

### ABSTRACT

The significant role of translation in the communication among individuals from various cultures and linguistic backgrounds has given rise to the pursuit of ways in which this communicative act might take place in a more rapid and fruitful manner. This was probably the main reason why Descartes, about a few centuries before now, endeavoured to build a common language system which all the people of our world could benefit. The idea of Descartes, which was to unite other languages in a system, has formed the main principle of the machines which were to carry out the coding and decoding functions and these machines were mainly developed for the purposes of extracting military intelligence from the foes. Here the main objective was to transfer the information coded in a text written in a given language into another language by stripping off the codes. This idea formed the very basis of Machine Translation (MT). From this point on, it was in some way aimed to automate the translation process by reducing the human translator's interference in the task. Yet, when it was understood that the human intervention was necessary in order to obtain translations of acceptable quality, we witnessed a shift from quest in the automation of translation process to the search for finding ways to facilitate the process by aiding human translators. Thus, the concept of machine translation has served to the emergence of a new notion by the name of Computer Assisted Translation (CAT).

This paper aims to describe the stages that the concept of machine translation has passed through beginning from the emergence of it and upto present day by pointing out the major historical milestones. Along these milestones, the events such as the first patents and ENIAC, Warren Weaver's vision, publication of the ALPAC report and the emergence of first commercial machine translation systems could be considered as having significant importance. As for the stages that the MT has passed through, reviving of systems used in machine translation such as Rule Based, Example Based and Statistical MT systems could be given.

Also there have been some hybrid MT systems and approaches in which more than one system are implemented. Nowadays, quests of combining artificial intelligence systems and MT systems have become a current issue and there is a growing interest in this area.

In the present study, factors contributing the development of MT have been touched upon along with the setbacks experienced in the practices of MT. Also areas where MT applications are commonly exercised are presented with some examples. And finally, remarks on which direction the MT practices might be heading is addressed.

**Keywords:** Machine Translation, Statistical Machine Translation, Rule-Based Machine Translation, Sample-Based Machine Translation

## 1. GİRİŞ

Öncelikle İngilizce terminolojide “Machine Translation” olarak adlandırılan ve Türkçe’ye genelde Makine Çevirisi olarak aktarılan kavramın aslında ne denli kabul edilebilir olduğu ya da bu vesileyle gerçekleştirilen eylemi ne kadar sağlıklı biçimde ifade ettiği tartışılabilir. Terimin kendisi Somers’in (Somers, 2003) de belirttiği üzere alan hakkında bilgisi olsun ya da olmasın birçok kişiye göre yanıltıcı bir anlam içerebilir. Öyle ki, ortada çeviri işlemini gerçekleştiren bir makineden ziyade bir bilgisayar bulunmaktadır ve bu bakımdan Somers “translation technology” (çeviri teknolojisi) ya da ‘*translation software*’ (çeviri yazılımı) gibi terimlerin daha kullanışlı olabileceği görüşünü paylaşmaktadır. Çünkü aslında makine çevirisinden bahsederken bilgisayarın donanımdan bileşenlerinden değil de yazılım bileşenleri kastedilmektedir. Türkçe’de ise benzer bir biçimde makine çevirisi ifadesiyle sıkça karşılaşılmasının yanı sıra bazı akademik metinlerde “bilgisayar çevirisi” ifadesiyle de karşılaşmak mümkündür (Şahin, 2015).

Terimin temsil ettiği kavramdan anlaşılması beklenen ise çeviri eylemini minimum düzeyde insan çevirmen müdahalesi ile (ki başlarda sıfır insan müdahalesinin hayali kurulmaktaydı) gerçekleştirmek üzere bilgisayarın ya da bilgisayar teknolojisinin önünü açtığı diğer teknolojik cihazların (cep telefonu, tablet bilgisayarlar, vb. gibi) kullanımınıdır. Ancak yine de Somers’in de açıkça belirttiği üzere (a.g.e) çeviribilim bütüncesinde henüz daha iyi bir alternatif terimin ortaya çıkmadığından bu çalışma kapsamında genel bakış açısından ayrılmamak adına makine çevirisi terimine bağlı kalınmıştır.

Makine çevirisi kavramı üzerine kısaca değindikten sonra mevcut çalışmanın temel amacı olan makine çevirisinin kısa tarihçesinin nasıl bölümlendirildiğine değinebiliriz. Makine çevirisinin tarihçesi üzerine farklı kaynaklar farklı olayları başlangıç noktası ya da belli başlı dönüm noktaları olarak ele almaktadır. Her ne kadar makine çevirisi fikrinin temellerinin onyedinci yüzyılda Rene Descartes’in evrensel bir dil oluşturma çabalarına dek eskilere dayandığı üzerinde genel bir uzlaşım olsa da (Ulitkin, 2011), bu fikrin uygulamaya geçirilmesinin ve zamanla ortaya çıkan yeni sistemler ile yaklaşımların tarihlerini belirlemede farklılıklar gözlemlenmektedir. Koehn’in istatistiksel makine çevirisi üzerine hazırladığı ayrıntılı çalışmada (2010, s. 14) makine çevirisi sistemlerini kurma çabalarının ilk bilgisayarların ortaya çıkmasının hemen ardından başladığı dile getirilirken John Hutchins (2004, s. 12) makine çevirisinin gerçek anlamda uygulamaya geçirilmesi üzerine Fransa ve Rusya’da eş zamanlı olarak iki farklı öncü tarafından elde edilen patentlere işaret etmektedir.

Ayrıca makine çevirisi ve genel tarihçesi üzerine yapılan çalışmaların çoğunda meşhur ALPAC raporu ortak bir yaklaşım ile bir dönüm noktası olarak belirlenirken bazı çalışmalar yirminci yüzyılın ortasından başlayarak hemen her on yıllık dönemi farklı birer dönüm noktası olarak ele almıştır (Qun & Xiaojun, 2015). Diğer yandan makine çevirisi üzerine araştırmalar ile uygulamaların tarihçesi üzerine hazırladığı çalışmada Hutchins 1950’lerden başlayarak 70’lere kadar olan dönemi başlangıç noktası sayıp, 70’lerden 1989’a kadar bölümü ikinci evre olarak adlandırmaktadır (2015, s. 120-136). Dahası, Hutchins 1990 yılından sonraki dönemi de sözlü dil çevirileri, çeviri araçlarının gelişimi ve derlem tabanlı araştırmalar gibi farklı başlık altındaki evrelere bölümlemiştir (a.g.e.).

Böylece makine çevirisi gibi aslında görece kısa bir geçmişe sahip bir kavramın tarihsel açıdan haritasının çıkarılmasında bile tam anlamıyla kesin bir uzlaşımın bulunmadığı görülmektedir. Bu sebeple mevcut çalışma kapsamında farklı kaynaklardan makine çevirisinin tarihçesinde ön plana çıkan önemli olaylar ele alınmış olup makine çevirisi kavramı için dönüm noktaları olarak kabul edilebilecek olaylar ya da uygulamalara dikkat çekilmiştir

## 2. EVRENSEL DİL ARAYIŞI

Rönesansın etkilerinin hemen her alanda tam anlamıyla hissedildiği on yedinci yüzyılda dönemin düşünürlerinin neredeyse tümü mantık ve deneye dayalı bilimin sağlam temellerini atma uğraşı içindeydiler. Bu doğrultuda dilbilim alanında çalışmalarını sürdüren düşünürlerin aklını kurcalayan üç temel konunun bulunduğu bahsedilir; evrensel dil, evrensel dilbilgisi ve dilin kökenleri (Jermolowicz, 2003). Aslında evrensel bir dil tasarlama düşüncesi yine Jermolowicz'in aktardığı üzere Milattan Önce 200 yılında Roma'da yaşamış olan bir doktor olan Claudius Galenus'a aitti (a.g.e). Ancak onun iletişim kurmayı kolaylaştırmak anlamına gelen uluslararası bir işaret sistemi kurma tasarısı gerçekleşmeden tarihin tozlu sayfalarında kaybolmuştu. Evrensel dil sisteminin temelinde de işte bu farklı diller kullanan insanlar arasındaki iletişimi kolaylaştırmak düşüncesi yatıyordu.

Galenus'tan yüzyıllar sonra, on üçüncü yüzyılda Katolik bir düşünür ve yazar olan Raymund Lullus terimlerin mantıksal bir sınıflandırılması ile evrensel dil arayışlarını sürdürmüş ve onun düşünceleri on yedinci yüzyıldaki düşünürlere ilham kaynağı olmuştur. On altıncı yüzyıla geldiğinde İtalyan bir şair olan Teofilo Folengo ise doğal bir dil üzerinden evrensel dil arayışına girmiş ve bir dizi İtalyanca lehçeyi kullanarak zenginleştirdiği Latincenin daha basit bir sürümü üzerinde çalışmıştır. Onun bu tasarısı daha sonraları ortaya çıkan Esperanto'nun da temelini oluşturmuştur. Hatta Thomas More'un önemli eseri Ütopya'da da Ütopyalıların kullandığı dil de benzer biçimde evrensellik özelliğine sahiptir.

On yedinci yüzyıla geldiğinde ise evrensel bir dil ihtiyacı daha da hissedilir olmuştur. Bunun en önemli nedeni olarak da gelişen ticaret yolları, misyonerlik faaliyetleri ile sömürgeciliğin artması olarak gösterilmektedir (Jermolowicz, 2003). Bu yüzyılda evrensel dillerin nümerik (sayısal) kodlar biçiminde oluşturulmasını tasarlayan düşünürler arasında Alman matematikçi Gottfried Leibnitz, Fransız düşünür René Descartes ve İngiliz yazar John Wilkins bulunmaktaydı (Somers, 2003, s. 4). Leibnitz'in evrensel dili "evrensel matematik" kavramına dayanıyordu ve dildeki işaretlerin matematikte olduğu gibi kesin bir biçimde kullanılması koşulunda düşüncenin matematikteki hesaplamalar gibi doğru biçimde anlaşılabilirliğini öne sürüyordu (Jermolowicz, a.g.e.). Descartes'a göre evrensel dil için atılması gereken ilk adım gerçekliği tamamen tanımlayan gerçek felsefenin keşfi idi ve bu yüzden de evrensel dil gerçek felsefenin doğal sonucu olabilirdi. Ancak o zaman bir dil düşünme sürecini yansıtabilirdi ve bu dilin oluşturulması da zihindeki temel düşüncelerin, düşüncelerin doğal sırasının keşfedilmesi ve her bir temel düşüncüyü bir sözcük ile birleştirmekten gibi basamakları izlemeliydi. Wilkins ise evrensel dil sistemi düşüncesinde yapay bir sembol sisteminin rolü üzerine odaklanmıştı. Bu yapay dil sistemi makine çevirisinde ileriki dönemlerde ön plana çıkacak olan '*ortak dil*' (interlingua) olarak adlandırılmaktadır (Hutchins & Somers, 1992). 1666'daki Büyük Londra Yangını'nda müsveddelerinin büyük bir kısmı yanan çalışmasında altı kategoriden oluşan bir terimler kümesi ileri sürmüştü ve daha sonra da bunları alt kategorilere bölmüştü. Böylece mantıksal bir harf dizisinin her bir kategoriyle eşleşmesinin mümkün olduğunu göstermiştir. Hatta bu sistemi kullanarak arkadaşı Robert Boyle ile yaptığı yazışmalarda düşüncesini uygulamaya da geçirdiği anlaşılmaktadır.

Sonuç olarak günümüzden yaklaşık olarak dört asır önce evrensel bir dil oluşturmaya yönelik bu ve benzeri girişimlerin makine çevirisinin temelini oluşturduğunu görmekteyiz. Her ne kadar bu girişimler tam anlamıyla uygulamaya geçirilememiş olsa da makine çevirisinin yirminci yüzyılın ortalarından itibaren başlayan ve günümüze dek uzanan uygulamaya yönelik serüveninin salt düşünsel bağlamda çıkış noktasına işaret etmektedir. Bu noktada, makine çevirisinin uygulamaya yönelik dönüm noktaları için ikinci bir başlangıç noktası belirlemek ve kısa tarihçesinin belli başlı önemli olayları şu şekilde belirlemek mümkündür; ilk patentler, ENIAC, Warren Weaver'ın öngörüsü, ALPAC raporu, ilk ticari sistemler, veri odaklı yöntemler, büyük oyuncuların hamleleri ve makinelerin eğitilmesi. Buradan hareketle de makine çevirisinin kısa tarihçesi üzerine belirli yıllar arasındaki dönemleri aşağıdaki sıralama içinde incelemek mümkündür.

## 3. MAKİNE ÇEVİRİSİNİN TARİHÇESİ

### 3.1. İlk Patentler ve ENIAC

Bilgisayarların henüz icat edilmediği bir dönemde, 1930'lu yıllarda, birbirinden bağımsız olarak çalışan Georges B. Artsrouni ile Petr Petrovič Trojanskij iki patent elde ettiler. Artsrouni'nin patenti 22 Temmuz 1933 tarihinde 'mekanik beyin' (mechanical brain) olarak adlandırdığı genel amaçlı bir cihaz içindi.

Cihaz genel amaçlı bir depolama işlevi görmekteydi ve demir yolu çizelgeleri, telefon rehberleri ve ticari telgraf kodları vb. gibi dökümanlardaki verilerin otomatik olarak üretilmesine mümkün kılmaktaydı. Ayrıca şifreleme, iletilerdeki şifrelerin çözülmesi ve yeniden şifrelenmesi gibi işlevler için uygundu. Bu sebeple de dillerin çevrilmesi için kullanılabilecek bir cihaz olduğu iddia edilmekteydi (Hutchins, 2004, s. 12-13).

Diğer yandan Trojanskij'in yine 1933 tarihli patenti makine çevirisi için daha önemli bir cihaza aitti. Onun makinesi birbirine paralel kolonlar içinde altı dilde girdileri bulunan kapsamlı bir sözlüğü içeren geniş bir banttı oluşuyordu. Trojanskij, cihazıyla mekanik çeviri sürecinin üç önemli aşamasını önceden tasarlamıştı: ilk aşamada sadece kaynak dili bilen bir düzeltmen sözcükleri kök biçimleri ile sözdizimsel işlevlerine göre mantıksal analizini gerçekleştiriyordu. İkinci aşamada ise makine kaynak dildeki sözcüklerin kök biçimlerini ve işlevlerinin erek dildeki karşılıklarına dönüştürüyordu. Son olarak ise sadece erek dili bilen ikinci bir düzeltmen elde edilen bu verileri erek dildeki normal biçimlerine dönüştürüyordu (Hutchins & Somers, 1992).

Ancak buki patentin ait olduğu makine ya da cihazların elektronik bilgi işleme ve depolama işlevlerine henüz sahip olmadıklarını da belirtmek gerekir. İkinci Dünya savaşının neden olacağı gelişmelerden biri ilk nesil bilgisayarların ortaya çıkması idi (Alberts & Papp, 1997, s. 25). Iowa Üniversitesinde çalışan John Vincent Atanasoff ilk çalışan veri işleme sistemleri ile bilgisayar belleğini 1939 yılında icat etti. Onun izinden giden İngiliz matematikçi Alan Turing ise Nazilerin savaş kodlarını kırmak ve Hitler'in savaş planlarına erişmek üzere kullanılan ilk dijital bilgisayarı, 'Colossus'u geliştirdi. Yine İkinci Dünya savaşında topçu birliklerinin ve roketlerin hedeflerini bulmalarına yardımcı olmak üzere yüksek hızda karmaşık matematiksel hesaplamalara ihtiyaç duyulması sonucunda J. Presper Eckert ve John Mauchly ENIAC'ı (*Electronic Numerical Integrator And Calculator* - Elektronik Sayısal Entegreli Hesaplayıcı) icat ettiler. Yaklaşık 17.000 elektron tüpü içeren bu bilgisayar 30 ton ağırlığındaydı ve yaklaşık 170 metre karelik bir alanı kaplıyordu. İlk nesil olarak adlandırılan bu bilgisayarlar daha çok kod kırma ve matematiksel hesaplama işlevleri için kullanılıyordu. Bilgisayarların çeviri eylemiyle tam anlamıyla bütünleştirilmesi düşüncesi ise savaşın ardından ortaya çıkacaktı.

### 3.2. Warren Weaver'in Öngörüsü

Bilgisayarın kod kırma işlevinden yola çıkarak çevirinin de en azından hesaplama bakış açısıyla kod kırma işlemine benzer olabileceği düşüncesi ilk olarak Warren Weaver'den çıktığı söylenebilir (Somers, 2003). Warren Weaver o dönemde Rockefeller Vakfının üst düzey yönetimindeydi ve "Çeviri" başlıklı bir bildirisinde MÇ'ye dair şu önerileri dile getirdi;

- Çoklu anlam sorunu birincil bağlamların incelenmesiyle çözülebilir
- Bütün dillerde ortak olan mantıksal özellikler bulunabilir
- İletişimin temel istatistiksel özellikleriyle ilgili şifreleme yöntemleri makine çevirisinde uygulanabilir
- Dilsel tümeller bulunabilir (Qun & Xiaojun, 2015, s. 105-106).

Bu önermelerin her biri ileriki dönemlerde MÇ üzerine çalışmalar için çok önemli bir yere sahip olacaktı. Kendisinden sonra gelenler tarafından kısmen ya da tamamen de olsa onun bu önerileri uygulamaya geçirilecekti.

1950'lerin başında MIT'de (Massachusetts Institute of Technology) tam zamanlı bir araştırmacı olarak göreve gelen Yehoshua Bar-Hillel ilk makine çevirisi konferansını 1952 yılında düzenledi ve bu etkinlik MÇ'ye ilgi duyanları bir araya getirip geleceğe yönelik çalışmalara yön verdi. Bundan iki yıl sonra IBM ve Georgetown Üniversitesi çalışanlarının ortaklaşa çalışmalarının ürünü olan Rusça'dan İngilizce'ye gerçekleştirilen ilk MÇ tanıtım gösterimi gerçekleştirildi. MÇ artık Birleşik Devletlerde milyon dolarlık bir iş alanı haline aldı ve 1960'ların ortalarına varmadan Avrupa kıtasından çoğu ülkenin yanı sıra Çin, Meksika ve Japonya gibi uzak coğrafyalarda da ilgiyle karşılandı. Ancak ne yazık ki araştırmaların çoğu MÇ'nin elde edilmesi güç ve yersiz Tam Otomatik Yüksek Kaliteli 'FAHQ' (Fully Automated High Quality) çeviriye ulaştırılmasını hedeflemekteydi. Elektronik beyinler olarak adlandırılan bilgisayarlardan beklentiler yüksekti ve kısa bir süre içinde ortaya çıkacak olan meşhur rapor da bir bakıma araştırmacıların ayaklarını yere basmalarında önemli bir etken olacaktı.

### 3.3. ALPAC Raporu

Bar-Hillel'in 1960'ların başlarında *Advances in Computers* (Bilgisayarlardaki Gelişmeler) başlıklı dergide yayınladığı çalışmasında Tam Otomatik Yüksek Kaliteli çevirileri hedefleyen MÇ araştırmacıları eleştirilmiş ve MÇ araştırmalarının görece daha az iddialı hedeflere odaklanıp insan ile makineler arasındaki etkileşimin daha ucuz yollarını araması gerektiğinden bahsedilmiştir (Qun & Xiaojun, 2015, s. 106). Bu yorumların MÇ çevreleri dışında pek fark edilmediği ve onun iddialarının da meşhur ALPAC raporu ortaya çıkana dek pek kabul görmediği anlaşılmaktadır.

Birleşik Devletler MÇ araştırmalarında en başından beri önemli bir rol oynamış ve bu alandaki çalışmalara maddi destek sağlamak konusunda en cömert uluslardan biri olmuştur. Ancak hali hazırda bilgisayarların Tam Otomatik Yüksek Kaliteli bir çeviriyi üretememiş olması bazı kuşkuların ortaya çıkmasına engel olamamıştı. Ulusal Bilim Akademisi 1964 yılında bir komisyon oluşturulup MÇ çalışmalarının beklentilerinin incelenmesine karar vermiş ve oluşturulan komite de 'Otomatik Dil İşleme Süreci Danışma Komisyonu' (Automatic Language Processing Advisory Committee: ALPAC) olarak adlandırılmıştır. Komisyon iki yıllık bir incelemenin ardından meşhur raporunu yayınladı. Aslında raporun başlığı 'Language and Machines: Computers in Translation and Linguistics' (ALPAC, 1966) yani 'Dil ve Makineler: Çeviri ve Dilbilimde Bilgisayarlar' biçimindedir ancak çeviribilimde ve özellikle de makine çevirisi araştırmalarında ALPAC raporu olarak tanınmaktadır. Raporda MÇ'nin ürünü olan çevirilerin insan çevirilerinden daha hızlı ya da ucuz olmadığına vurgu yapılmakta ve kısa vadede MÇ'nin uygulanabilirliği ya da yararlı olabileceğine dair herhangi bir umudun olmadığı dile getirilmektedir. Ayrıca bu alana yapılan yatırımların da MÇ sistemlerini kullanmanın herhangi bir olumlu yanı bulunmadığı için insan çevirilerini geliştirme yöntemleri ile temel dilbilim araştırmalarına yöneltilmesinin gerekliliğinden bahsediliyordu. Bu doğrultuda ise otomatik sözlükler gibi insan çevirilerine yardımcı olabilecek araçların geliştirilmesi tavsiye ediliyordu (Hutchins, 2015, s. 122).

Raporun özellikle Birleşik Devletler'deki MÇ araştırmaları üzerinde olumsuz etkileri oldu ve neredeyse bu alandaki çalışmalar durma noktasına geldi. Bu arada alana yapılan büyük ölçekli ve devlet destekli mali yardımlar da kesintiye uğradı. Raporun bu denli olumsuz yargılar ile sonuçlanması MÇ araştırmalarıyla ilgilenen çevrelerce raporun dar görüşlü, önyargılı ve öngöründen uzak olduğu biçiminde yorumlanmasına neden oldu. Ancak yine de MÇ araştırmaları Birleşik Devletler dışında özellikle de Kanada, Fransa ve Almanya gibi ülkelerde ilerlemeye devam etti.

### 3.4. Yeniden Diriliş

ALPAC raporunun olumsuz etkilerinin Birleşik Devletler'deki kadar şiddetli bir biçimde hissedilmediği Kanada'nın Montreal kentinde başlatılan TAUM (Traduction Automatique à l'Université de Montréal) projesi kapsamında iki önemli sistem geliştirildi; Q-systems (Q-sistemleri) ve Météo. İlki dilsel dizilerin ve sözdizim ağaçlarının kullanımına dayanıyordu. İkincisi ise hava durumu tahminlerinin çevirisine odaklanmıştı. Hava durumu tahminlerinin görece dar bir kelime dağarcığına ve sınırlı bir sözdizim çeşitliliğine sahip olmalarının sonucu olarak Météo sistemi halen başarılı bir biçimde kullanılmaktadır. TAUM grubunun benzer bir başarıyı hedefleyerek giriştikleri havacılık ile ilgili kılavuzların çevirileri süreci ise karmaşık bileşik isim ve tümce yapılarından dolayı ise istenilen başarıyı yakalayamamıştır.

Diğer yandan Fransa'da, 1960 ile 1971 arasında, Grenoble Üniversitesinde Bernard Vauquois'nun öncülüğünde oluşturulan bir grup olan CETA (*Centre d'Etudes pour la Traduction Automatique*) ya da diğer bir adıyla GETA (*Groupe d'Etudes pour la Traduction Automatique*) matematik ve fizik alanlarından metinleri Rusça'dan Fransızca'ya çevirmek üzere *interlingua* (ortak dil ya da *pivot dil*) sistemini geliştirdi. Ancak sistem sözdizimsel bağıntıların mantıksal özelliklerinin temsil edilmesi üzerine dayanıyordu ve sözcüklerin diller arası anlatımlarını sağlamadığı için tam bir ortak dil sayılmazdı. Onun yerine sözcükler çift dilli bir aktarım mekanizması aracılığı ile çevriliyordu. Sözdizimsel analiz sayesinde bir tümce-yapısı simgelemi oluşturuluyor ve ardından bağımlılık ilişkileri ekleniyor ve sonunda özne ve yüklemelere göre bir ortak dil simgelemesi üretiliyordu (Hutchins, 2015, s. 124). Benzer bir model Texas Üniversitesinde 1970'li yıllarda Dilbilim Araştırma Merkezi (LRC – The Linguistics Research Center) tarafından geliştirilen METAL sisteminde kullanıldı.

Yine 1960'ların sonlarına doğru Dr. Peter Toma tarafından kurulan SYSTRAN sistemi Birleşik Devletler Hava Kuvvetleri tarafından 1970'ten itibaren kullanılmaya başladı. İlk başta Rusça-İngilizce

dil çiftinde kullanılan sistemin Avrupa Komisyonu tarafından Fransızca-İngilizce dil çifti sürümü satın alındı. Bu gelişme ayrıca daha çok sayıda Avrupa dili çifti için sistemlerin gelişmesine de ön ayak oldu. Systran sistemi ayrıca Xerox firması tarafından da 1970'lerin sonlarından beri kullanılmaktadır. Dilsel öğelerin doğrudan aktarılmasına dayanan bu gibi sistemler genel olarak doğrudan çeviri (*direct translation*) yaklaşımını temsil edip birinci nesil makine çevirisi sistemleri olarak adlandırılabilir. *Doğrudan çeviri* yaklaşımını temel alan bir başka sistem de Logos sistemi idi. Sistem 1970'lerde savaş uçakları için hazırlanan kılavuz metinlerinin İngilizce-Vietnamca dil çiftindeki çevirisinde kullanılmış olan bir başka önemli sistemdi.

Ancak ortak dil sistemlerinin hayal kırıklığına neden olmasının ardından GETA ikinci nesil dilbilim tabanlı aktarım sistemlerinin örneğini teşkil edecek olan Ariane sistemini geliştirdi. Deneysel sistemler arasında sayılan Ariane, 1980'lerde dünya çapında birçok proje üzerinde etki yaratmıştır. Bunlar arasında Kyoto Üniversitesinde Makoto Nagao tarafından geliştirilen *Mu* sistemi ve Almanya Saarbrücken'de geliştirilen *SUSY* (*Saarbrücker Übersetzungssystem – Saarbrück Çeviri Sistemi*) sayılabilir. *SUSY*, çokdilli bir aktarım sistemiydi ve tümce yapısı kuralları, dönüşümsel kurallar, durum dilbilgisi, bağımsal dilbilgisi ve istatistiksel veri kullanımı gibi birbirine benzemeyen birçok tekniği içeriyordu.

1980'lere gelindiğinde ise Avrupa Komisyonu'nun kullandığı SYSTRAN sisteminin '*doğrudan çeviri*' yaklaşımının aslında sınırlı olduğu varsayımından yola çıkarak hedeflenen Eurotra projesi dönemin en önemli makine çevirisi sistemlerinden biri oldu. Dil dışı bilgi tabanlarının ya da etki mekanizmalarının kullanımını gerekmeyordu. Ayrıca çeviri sürecinde de herhangi bir insan yardımı ya da müdahalesi içermiyordu. Ancak sistemin en önemli kusuru sözcük düzeyinde ortaya çıkan sorunların çözülmesinde karşılaşılan başarısızlık idi.

1980'lerin ortasından itibaren yapay zekâ ile bilişsel dilbilim alanlarında ortaya çıkan araştırmalar ortak dil (*interlingua*) sistemlerinin yeniden ilgi odağı haline gelmesine neden oldu. Toon Witkam'n yönetimi altında Hollanda, Utrecht'te başlatılan DLT (Distributed Language Translation – Dağıtılmış Dil Çevirisi) sistemi Esperanto'nun geliştirilmiş biçiminde oluşturulan bir ara dilde birbirine bağlı terminaller arasında metinlerin aktarılmasını içeriyordu. Benzer biçimde ortak dil kullanan bir başka sistem yine Hollanda'da Eindhoven'da başlatıldı: Rosetta projesi.

Hem özel hem de kamu sektörü tarafından yaklaşmakta olan yeni nesil bilgi toplumu için gerekli olarak nitelendirilen MÇ araştırmalarına büyük yatırımların yapıldığı 1980'lerin ikinci yarısına gelindiğinde Fujitsu, Toshiba ve Hitachi gibi Japon şirketlerinin ön plana çıktığı görülmektedir (Hutchins, 2015, s. 125). Özellikle mikrobilgisayarların ve metin işleme sistemlerinin artık daha yaygın bir biçimde erişilebilir olması daha ucuz MÇ sistemleri için önemli bir pazarın ortaya çıkmasını sağlamıştır. Yine aynı dönemde görece daha az yer kaplamasına karşın sözlü dil çevirisi için MÇ sistemleri için araştırmaların başladığını görmekteyiz. Bu sistemler başlangıçta dil kılavuzları ve bilgisayara tabanlı sistemler olarak ortaya çıkmıştır. Bu alanda Globalink'in 1995 yılında yazılı metinden ses çıktısı alınmasını sağlayan sistemleri ile öncülük ettiği söylenebilir.

1970'ler ve 1980'lerde ortaya çıkan sistemlerde ön plana çıkan doğrudan aktarım (transfer) ve ara dil (*interlingua*) yaklaşımlarına ek olarak derlem tabanlı (*corpus-based*) yaklaşımını da MÇ alanındaki araştırmalar için önemli bir dönüm noktasıdır. Önceki sistemlerde kullanılan yaklaşımlar daha çok kural tabanlı olarak adlandırılabilir: sözdizim analiz kuralları, sözdizimsel kurallar, sözcük aktarım kuralları, sözdizim oluşturma kuralları ya da biçimbilim kuralları gibi (Hutchins, a.g.e.).

Kural tabanlı MÇ yaklaşımından derlem tabanlı yaklaşımlara geçiş döneminde yaşanan en önemli gelişme tamamen istatistik tabanlı yaklaşımın geri dönüşü oldu. IBM şirketi çalışanları tarafından İngilizce – Fransızca dil çifti için geliştirilen Candide projesi aslında uzun zaman önce Warren Weaver'in öngördüğü istatistiksel yaklaşıma dayanıyordu. Ancak Candide projesinin temelinde çift dilli bir derlemeden otomatik olarak elde edilen istatistiksel bilgi ile insan uzmanların dil bilgisini olasılıksal bir çerçeve içinde birleştirme düşüncesi yatıyordu (Pietra & Pietra, 1994). Sistem Kanada parlamentosunda gerçekleşen müzakerelerin raporlarından alınan Fransızca ve İngilizce metinlerin oluşturduğu iki büyük derlem üzerinde test edildi ve sonuçların kabul edilebilir düzeyde olduğu ortaya çıktı.

Makine çevirisi çalışmalarının istatistiksel sistem üzerinden yürütülmesi bazı araştırmacılar tarafından yeniden '*doğrudan aktarım*' yaklaşımının '*yeniden dirilişi*' olarak adlandırılmaktadır; bkz. Hutchins

(2015, s. 128) ve Qun & Xiaojun (2015, s. 107). Bu bağlamda ele alınan istatistiksel makine çevirisi sistemleri önce çift dilli derlemlerin birbirleri ile *hizalanması* (alignment) daha sonra da girdi sözcüklerinin derlemlerdeki sözcükler ile eşleşme frekanslarının bulunup erek dildeki en olası eşdeğerliklerinin ortaya çıkarılması '*decoding*' yani şifrelerinin çözülmesi işlemi, erek dilin sözcük frekanslarını sağlayan tek dilli bir derlem olarak adlandırılabilen bir 'dil modeli' kullanarak en yaygın sözcük dizilimlerine göre çıktılarının yeniden düzenlenmesi ve son olarak da erek dilde çıktının üretilmesi gibi basamakları içermektedir.

İstatistiksel MÇ sistemlerinde önceleri *sözcük tabanlı* (word based) olarak ortaya çıkan bu yaklaşımın daha sonra *tümce tabanlı* (phrase based) ve *cümle tabanlı* (sentence based) olarak türevlerinin ortaya çıktığından söz etmek mümkündür. Şöyle ki, ilk başlarda kaynak ve erek diller arasındaki sözcük bağıntılarına (correlation) odaklanan sistemlerde sonraları bunun yerine sözcük dizileri, bir başka deyişle tümceler arasındaki bağıntılar üzerinde durulmuş ve bu arada da sözlük gibi kaynaklar kullanılarak söz konusu diller ile ilgili morfolojik (biçimbilimsel) ve sentaks (sözdizimsel) bilgiler dâhil edilmiştir. Bu durum *kural tabanlı* (rule-based) yaklaşımlardan artık *derlem tabanlı* yaklaşımlara doğru bir yönelimin başladığının da işareti olmuştur.

Her ne kadar Nagao tarafından '80'lerde ilk kez öne sürülmüş olsa da derlem tabanlı sistemlerde önemli bir dönüm noktası da *örnek tabanlı* (example based) yaklaşım olmuştur (Hutchins, 2015, s. 129). Örnek tabanlı yaklaşımın temelinde insan çevirmenlerin ürettikleri çevirilerin çoğunlukla benzeşik (analogous) örneklerin bulunması ya da belirli bir kaynaktan geri toplanması (recall) işlevini içerdiği varsayımı bulunmaktadır. Burada kaynak olarak nitelendirilen aslında daha önceden gerçekleştirilmiş çevirileri temsil etmektedir ve çeviri sürecinde bir eşleştirmeden yararlanılmaktadır. Bu tip bir eşleştirmeden yararlanılması günümüzde çeviri çeviri bellekleri olarak bilinen bilgisayar destekli çeviri araçlarından birinin yaygınlaşmasına da katkıda bulunmuştur. Ayrıca Tak-ming ve Webster'in de ifade ettiği üzere (2015, p. 146) örnek tabanlı yaklaşım gerçek dil kullanımını temsil etmektedir ve gerçek anlamda daha önceden oluşturulmuş dil yapılarını kapsamaktadır ve kolayca yeni örnekler eklenerek genişletilmesi mümkündür. Ancak daha önceden çevirilmemiş tümce ya da cümlelerin çevirisinde kullanımı kısıtlı olup insan çevirmenler ya da düzeltmenlerin müdahalesini gerektirmektedir.

### 3.5. Sözlü Dil Çeviri Sistemleri

Makine çevirisi bağlamında sözlü dilin çevirisi olgusu metin-söz, söz-metin ya da söz-söz eşleşmeleri doğrultusunda ele alınabilir. Bunlardan ilki herhangi bir dilde yazılı olarak üretilmiş bir metnin başka bir dile sözlü olarak aktarılması, ikincisi tam tersine sözlü olarak üretilmiş bir kaynak metnin erek dile yazılı olarak aktarılması ve üçüncüsü ise sözlü olarak üretilen kaynak metnin erek dile yine sözlü olarak aktarılması biçiminde yorumlanabilir. Bu alanda ilk çalışmaların 1988 yılında Japonya'da ortaya çıktığını görmekteyiz (Hutchins, 2015, s. 131). Japonya'da ATR (Advanced Telecommunication Research – Gelişmiş Haberleşme Araştırması) araştırma kurumunun başlattığı çalışmaları Amerika'da Carnegie Mellon üniversitesindeki JANUS sistemi ile Almanya'daki Verbmobil sistemi 1990'larda takip etmiştir. Yine aynı dönemde Cambridge'teki SRI grubunun başlattığı SLT (Spoken Language Translator – Konuşma Dili Çeviricisi) projesi de konuşma dili çevirisi üzerine bir sisteme odaklanmıştır. 2000'li yılların başında da Birleşik Devletler Ulusal Bilim Vakfı ile Avrupa Komisyonu'nun ortaklaşa desteklediği NESPOLE (NEgotiating through SPOken Language in E-commerce – Elektronik Ticarete Sözlü Dil ile Uzlaşım) araştırma projesi sözlü dil çeviri sistemleri arasında sayılabilir.

ATR araştırmalarının bir ürünü olarak Toshiyuku Takezawa ve ark. (1998) tarafından geliştirilen ATR-MATRIX (Multilingual Automatic Translation System for Information Exchange – Bilgi Alışverişi için Çokdilli Otomatik Çeviri Sistemi) sistemi bir ana kontrol ünitesine bağlı üç alt sistemden oluşmaktaydı. Ayrı ayrı olarak ana kontrol sistemine uydu bağlantısı ile eklenen bu üç alt sistem ise konuşma metinlerinin çözümlenmesi sağlayan bir ses tanıma alt sistemi, tanımlanan dil girdilerinin çevirilerinin gerçekleştiği bir dil çevirisi alt sistemi ile elde edilen dil verilerini sesli ortama aktaran ses sentezleme alt sisteminden oluşmaktaydı. Alt sistemlerin birbirleri ile doğrudan bağlantılı olmaması ana kontrol sisteminin her bir alt sistemle eşit biçimde etkileşimde bulunmasını sağlıyordu. Sistem ilk olarak otel odası rezervasyon görüşmelerinin gerçekleştirildiği bir bağlamda kullanıldı.

JANUS etkileşimli bir konuşma dili çevirisini amaçlayan geniş çaplı bir proje olmasına rağmen geliştirildiği 90'lı yıllarda İngilizce, Almanca ya da İspanyolca dillerinden sınırlı alanda

gerçekleştirilmiş konuşmalardan eşzamanlı olarak dil girdilerini alıp Almanca, İngilizce, İspanyolca, Japonca ve Korece’de çıktılar vermekte idi (Waibel, Lavie, & Levin, 1997). Sistem başlıca ses tanıma, söz dizimsel analiz, söylem oluşturma ve üretim aşamalarına sahip olup bir ara dil tabanlı yaklaşımı izlemektedir.

Almanya’nın yanı sıra Amerika ve Japonya’dan çeşitli kurum ve kuruluşların desteği ile başlatılan Verbmobil sistemi de benzer biçimde eş zamanlı ve kesintisiz konuşma metinlerinin sözlü olarak çevirisine odaklanmıştı (v. Hahn & Amtrup, 1996). Başlangıçta çeviri süreci için Almanca ve Japonca dil çifti seçilmiş ve İngilizce ise ara dil (temsili dil) olarak kullanılmıştır. Konuşmaların gerçekleştiği bağlam olarak randevuların planlandığı konuşmalar seçilmiştir. Daha sonraki proje çalışmalarına seyahat planlamaları üzerine gerçekleştirilen konuşmalar dâhil edilmiştir.

SRI grubunun geliştirdiği proje sayesinde ortaya çıkarılan DECIPHER (TM) sistemi İngilizce – İsveççe dil çiftinde gerçekleştirilen havayolu seyahatlerinin planlanmasına dair konuşmalar üzerinde başlatıldı. Sistemin altyapısının diğer sistemlere görece daha sağlam olduğu ve diğer diller ile farklı alanlara daha kolay taşınabileceği iddia ediliyordu (Rayner, ve diğerleri, 1993, s. 217-222). Buna neden olarak sistemin bileşenlerinin diğerlerine göre daha modüler olması gösterilebilir. Öyle ki, sistem içinde ses tanıma, kaynak dil işleme, kaynaktan erek dile aktarım, erek dil işleme ve son olarak da *konuşma biresimi* (speech synthesis) gibi farklı bileşenlerden oluşmaktaydı.

Son olarak NESPOLE araştırma projesinin ürünü olan sistem İtalyanca, Almanca, Fransızca ve İngilizce dilleri arasında çok dilli ve söz-söz temelli etkili bir çeviri işlemi için tasarlanmıştı (Lavie, ve diğerleri, 2001, s. 31). İlk olarak turizm ve seyahat bilgilendirme alanlarında kullanıldı. Sistemin etkili olarak elektronik ticaret ortamlarında kapsamlı ancak sınırlı bir alan yelpazesinde kullanılması hedeflenmekteydi. Araştırmacılar özellikle elektronik ticarete konuşma, video ve mimiklerin de dâhil olduğu çok kipli etkileşimin teknoloji tabanlı ticarete önemli bir rol oynayacağı görüşünden yola çıkarak tedarikçiler ile müşteriler arasındaki iletişimin çevrimiçi etkileşimi desteklemesinin gerektiğini ileri sürmektedirler (Lavie, ve diğerleri, a.g.e.). Bu öngörü teknolojinin hayatımızın hemen hemen her alanına girdiği günümüz koşulları için yeterince geçerli sayılabilir.

Benzer biçimde IBM’in DARPA CAST için geliştirdiği *MASTOR* (Multilingual Automatic Speech-to-Speech Translator – Çokdilli Otomatik Söz-Söz Çeviricisi) sistemi de mobil iletişim araçlarında geçek zamanlı söz-söz doğrultusunda çevirinin yaygınlaştırılmasını hedeflemektedir (Dureja & Gautam, 2015, s. 28). Ancak sözlü dil çevirisinin makineler ya da daha doğru bir ifade ile bilgisayarlar tarafından tamamen otomatik olarak gerçekleşmesinin önünde birtakım engellerin bulunduğu da bir gerçektir.

Sözlü dil çevirisinin tamamen makine çevirisi sistemleri tarafından gerçekleştirilmesini kısıtlayan faktörler insan sesine ait verilerin çeşitliliği, yazılı dilde kullanılan noktalama işaretlerinin sözlü dilde tonlamalar, duraksamalar ya da farklı mimik ve jestlerle ifade edilebilmesi, sözlü dilin genel dil normlarının dışındaki kullanımlara sürekli açık ve bağlama dayalı olması, kişilerarası gerçekleşip çoğunlukla diyaloglar biçiminde üretilmesi ve özellikle rakamlar ile sayısal ifadelerin yazılı dilden farklı bir biçimde kullanılması vb. olarak özetlenebilir. Her ne kadar sözlü dil çevirisi üzerine sınırlı alanlarda oldukça çok sayıda araştırma yapılmasına karşın hali hazırda ticari anlamda uygulanabilir tam otomatik bir sözlü dil makine çevirisi sistemi bulunmamaktadır.

### 3.6. Yeni Asır Yeni Yaklaşımlar

Yirminci yüzyılın sonları ile yirmibirinc yüzyılın başlarından itibaren hem akademik hem de ticari sistemlerin birçoğunda istatistiksel makine çevirisi sistemlerinin hızlı bir biçimde gelişimine tanıklık etmekteyiz. Qun & Xiaojun’un da ifade ettiği gibi (2015, s. 109) IBM şirketi bu dönemdeki araştırmalara öncülük etmiştir ve IBM Model 1-5’e dayanan sözcük tabanlı makine çevirisi yöntemlerini ileri sürmüştür. IBM’in bu modeli iğher tüm makine çevirisi sistemleri için temel olan sözcüklerin hizalanması temeline dayanıyordu. Sözcük tabanlı sistemlerden tümce tabanlı sistemlere doğru yönelim bir süre sonra sentaks (sözdizim) tabanlı modellerin de ortaya çıkmasına neden oldu. Qun & Xiaojun’un da işaret ettiği üzere tipik sentaks (sözdizim) tabanlı modeller arasında ters dönüşüm dilbilgisi, hiyerarşik tümce-tabanlı model, dizi-sözdizim ağacı ve sözdizim ağacı-dizi modelleri bulunmaktadır. Ayrıca büyük hacimli derlemlere erişilebilmesi sonucunda paralel hesaplamalar ve sinir ağları ve ağlar arasındaki bağıntılar üzerine yapılan çalışmalar ivme kazandı.



Romanov ve arkadaşlarının da belirttiği üzere modern dilbilim mühendisliğinin temel görevi Doğal Dil İşleme (DDİ: NLP – Natural Language Processing) ile MÇ sistemlerinin sinerjetik yeteneklerinin geliştirilmesidir (Romanov, ve diğerleri, 2003, s. 215-217). Bunu sağlamak için de bir dil otomasyonunun içerisine sözlü bir iletiyi üretme ve algılamada insanın sözlü ve mental etkinliğini kontrol eden iletişimsel-edimsel bir operatörün eklenmesi gerekir. Şu ana dek böyle bir durum söz konusu olmadı ya da bir başka deyişle böyle bir durumun sonucu olarak gerçek anlamda tam otomatik yüksek kalitede çeviri üretebilecek bir bilgisayar icat edilmedi.

Ancak ikinci nesil DDİ ve MÇ sistemleri geliştiricileri çalışmalarını iki önemli kuramsal temele dayandırdılar. Birincisi insan dilinin belirsiz nesnelere ile ilgilenen çağrışıma açık ve devingen bir sistem olarak ele alınması gerektiğini bu yüzden de belirsiz sayıda küme ve belirsiz bir mantıktan oluştuğu kuramı ile desteklenmesi gerektiği düşüncesidir. İkincisi ise her bir insan dilinin kendine has etnodilbilimsel özelliklere sahip olduğu ve bu özelliklerin de gerçek bir metnin ve onun bileşenlerinin dilici ve diller arası özelliklerinin kaynağı olduğu düşüncesidir (Romanov, a.g.e.). Bu düşüncelerin uygulamaya geçirilmesi sürecinde davranış-tabanlı yapay zekâ ve MÇ kavramları odak noktası olmaya başladı.

Diğer yandan ses tanıma ve resim ya da fotoğraf çevirisi gibi görsel verilerin içinde gizli olan metinlerin çevirilebilmesi gibi özelliklerin makine çevirisi sistemlerine dâhil edilmesi yeni bir istatistiksel çağa girildiğini göstermektedir. IBM'in sözcük tabanlı modelleri ile Gizli Markov Modeli'nin (HMM – Hidden Markov Model) uygulanması olarak görülen GIZA ++ günümüzde makine çevirisinde sözcük hizalamasında yaygın bir biçimde kullanılmaktadır. Diğer yandan Moses sistemi ise tümce tabanlı bir sistem olup makinenin ya da bilgisayarın eğitilmesini içermektedir. Bilgisayarların eğitilmesi kavramı yapay zekâ çalışmalarının makine çevirisi ile birleştirilmesi ile ilişkilendirilebilir.

MÇ'nin tarihçesinde son olarak değinilmesi gereken bir konu da açık kaynaklı makine çevirisi teknolojisidir. Genel olarak makine çevirisi yazılımları başlığı altında toplanabilen bu teknoloji ticari lisanslı makine çevirisi yazılımları ve ücretsiz/açık kaynaklı makine çevirisi yazılımları olarak sınıflandırılabilir (Forcada, 2015). Ticari makine çevirisi sistemlerinin çoğu kural tabanlı olup tamamen umuma açık olmayan tescilli teknolojilere sahip arama motorlarını içerirler. Sistemdeki dil verilerinin son kullanıcılar tarafından değiştirilmesine genellikle izin verilmez ancak yine de sistemin sözlük ya da terminoloji tabanlarına yeni veri girişi kullanıcılar tarafından gerçekleştirilebilir. Ayrıca kullanıcılar tarafından sistemin yeni bir dil çifti için sil baştan yapılandırılması da mümkün değildir.

Kural tabanlı bir makine çevirisi sisteminin ücretsiz ya da açık kaynaklı olması için belirli bir dil çifti için dil verilerinin kaynak kodunun yanı sıra arama motoru ve araçları için de kaynak kodlarının ücretsiz olarak ya da açık biçimde dağıtımının mümkün olması gerekmektedir. Böylece ücretsiz/açık kaynaklı makine çevirisi kullanıcıları dil verilerini değiştirebilirler. Şayet sistem derlem tabanlı bir yaklaşım üzerinden geliştirilmiş ise kaynak kodlarının cümle düzeyinde hizalaması yapılmış paralel metinler ile birlikte dağıtılmış olması gerekir.

Bazı makine çevirisi sistemleri ise ne tam anlamıyla ticari ne de tam olarak ücretsiz/açık kaynaklı makine çevirisi sınıflandırmasına girerler. Günümüzde genel ağ üzerinden erişilebilen Google Translate ya da Bing Translator gibi genel ağ tabanlı sistemler derlem tabanlı olup küçük hacimli metinler için ücretsiz ya da açık kaynaklı makine çevirisi sistemlerine örnek verilebilir. Google Translate ve Bing Translator derlem tabanlı modeller olup Bunun dışında, örneğin Türkçe-İngilizce için, Tureng ve seslisözlük gibi bazı ağ tabanlı sözlük sistemleri de kullanıcıların veri girişine izin vermektedir.

Son olarak özellikle bilgisayarların eğitilmesini temel alan derin öğrenme kavramının makine çevirisi sistemlerine de uyarlanmaya başladığına tanıklık etmekteyiz. Makine çevirilerinin değerlendirilmesinde de kullanılan BLEU (Bilingual Evaluation Understudy – Çiftilli Değerlendirme Dublörülüğü) yöntemi derlem veritabanı kullanarak gerçekleştirilen çevirilere 0 – 1 arasında bir skor değeri vererek karşılaştırma gerçekleştirir. Bu değerlendirmeyi yaparken kullanılan referans daha önce gerçekleştirilmiş ve görece daha yüksek bir skora sahip olan çevirilerdir. Böylece sistemin ileride benzer kaynak metinlerle karşılaşması durumunda görece yüksek skorları olan çevirileri kullanması sağlanır. Ancak herhangi bir kaynak metin cümlesinin bile 'mükemmel' birçok çevirisi olabilir (Papineni, Roukos, Ward, & Zhu, 2002). Ayrıca bu çeviriler sözcük seçiminde farklılık gösterebileceği gibi aynı sözcüklerin kullanılması durumunda da sözcüklerin dizilimi konusunda çeşitlilik gösterirler. İnsanların başarılı

çevirileri görece başarısız olanlardan ayırabilmeleri mümkündür ancak çevirilerin sadece insanlar tarafından değerlendirilmesi bazı durumlarda uzun zaman alabilir ve haliyle pahalıya mal olacaktır. Buradan hareketle geliştirilen yeni MÇ sistemleri bilgisayarların eğitilerek çeviri değerlendirmesi sürecini kolaylaştırmayı hedeflemektedirler. Bunun da ötesinde nöral makine çevirisi olarak adlandırılan yeni yaklaşım ile makinelerin büyük çapta yapay bir sinir ağı kullanması üzerine çalışmalar yürütülmektedir. Makine, bu sinir ağı ile sisteme dâhil edilmiş tek bir modeldeki bütün cümleleri örnek alarak bir sözcük dizisinin çevirisi üzerine tahminlerde bulunmaktadır. Bu alandaki araştırmaların görece yeni olduğunu göz önünde bulunduracak olursak henüz başarısı hakkında herhangi bir yorum yapmak şimdilik mümkün değildir. Ancak yine de bu yeni yaklaşımın makine çevirisinin tarihçesinde ilk dönemlere damgasını vuran tam otomatik yüksek kaliteli makine çevirisi hayalinin yeniden canlanmış olabileceği düşünülebilir.

#### 4. SONUÇ

Görece kısa bir süre önce girmiş olduğumuz yeni binyılda şu ana dek tanık olduğumuz teknolojik gelişmeler ister istemez birçoğumuzun bilim kurgu filmlerinde gördüğümüz sahnelerin birer birer gerçeğe dönüşebileceğine dair inancını kuvvetlendirmektedir. Özellikle yapay zekâ üzerine yürütülen çalışmaların çeviri teknolojileri ile birleştirilmesine yönelik girişimlerin bazı çevrelerde sanki makinelerin ya da bilgisayarların çevirmenlerin yerini alabileceği biçiminde yorumlanmasına da tanıklık etmekteyiz. Ancak özellikle yukarıda sözlü dil çevirisi sistemleri ile ilgili olarak değinilen sorunların yanı sıra makine çevirisinin yazın çevirisi gibi yoruma açık, belirli bir terminolojiye sahip olmayan, öznel anlatımların sıkça karşılaşıldığı koşullarda yetersiz kalabileceği kaçınılmaz bir gerçektir. İnsan çevirmenlere olan ihtiyaç yadsınamaz. Huthcins ve Somers'in makine çevirisi üzerine yaptıkları çalışmanın giriş bölümünde de vurguladıkları gibi; '*insan müdahalaesi ya da yardımı olmaksızın herhangi bir dildeki cümleyi alıp birkaç tuşa basarak bir başka dile mükemmel derecede çevirisini yapabilecek bir çeviri makinesi yoktur*' (1992, s. 1). Tamamen otomatik yüksek kalitede çevirinin gerçekleşmesi güç bir hayal olduğu ancak aşırı tekrarın bulunduğu, görece sınırlı terminolojiye sahip ve yoruma açık olmayan metinlerin çevirisinde makine çevirisinin belirli bir ölçüde başarıya ulaşabileceğinden söz etmek mümkündür.

Yine de makine çevirisinin görece kısa tarihçesi içinde önemli adımların atıldığından da bahsetmek mümkündür. Başlangıçta kod çözme işlevi gören ve kabaca elektronik sözlük görünümlü basit makinelerden günümüzde büyük çapta veriyi kısa sürede analiz edebilen ve birden çok dile çeviriyi mümkün kılan karmaşık sistemlere ulaşmış durumdayız. Hatta bulut teknolojisinin yaygınlaşması sonucunda genel ağ üzerinde çevirileri saklama ve daha önceden saklanmış çeviri verilerine ulaşma gibi görevleri de birkaç tuşa basarak yerine getirmek mümkündür. MÇ'nin günümüze dek süren yolculuğunda bilgisayarlar ile paralel bir doğrultuda gelişim gösterdiğini görmekteyiz. Bilişim teknolojilerinde gerçekleşen birçok yeniliğin eş zamanlı olmasa da görece kısa bir süre içinde makine çevirisi sistemlerine de uyarlandığı bir gerçektir. Çeviri piyasasında hali hazırda kullanılan birçok sistemin de bu gelişmelerden yararlandığı gerçektir. Görece makine çevirisi kavramından daha sonra ortaya çıkan bilgisayar destekli çeviri (BDÇ) araçları da bu gelişmelerin bir sonucudur. Bu bağlamda teknoloji ile çevirinin bir araya gelmesi hem MÇ hem de BDÇ sistemlerinin aracılığıyla gerçekleşmektedir. Ancak her iki durumda da insan faktörünün göz ardı edilmesi şimdilik mümkün görünmemektedir.

#### KAYNAKÇA

Alberts, D., & Papp, D. (1997). *The Information Age: An Anthology on Its Impact and Consequences*. Washington DC.: National Defense University Press.

ALPAC. (1966). *Language and Machines: Computers in Translation and Linguistics*. Washington, D.C.: National Academy of Sciences & National Research Council. Ekim 3, 2018 tarihinde [https://www.nap.edu/resource/alpac\\_lm/ARC000005.pdf](https://www.nap.edu/resource/alpac_lm/ARC000005.pdf) adresinden alındı

Dureja, M., & Gautam, S. (2015). Speech-to-Speech Translation: A Review. *International Journal of Computer Applications*, 129(13), 28-30. Kasım 15, 2018 tarihinde <https://www.ijcaonline.org/research/volume129/number13/dureja-2015-ijca-907079.pdf> adresinden alındı

- Forcada. (2015). Mikel L. C. Sin-wai (Dü.) içinde, *The Routledge Encyclopedia of Translation Technology* (s. 152-166). Londra & New York: Routledge Taylor & Francis Group.
- Hutchins, W. (2004). Two Precursors of Machine Translation: Artsrouni and Trojanskij. *International Journal of Translation*, 16(1 (Jan-Jun)), 11-31. Eylül 24, 2018 tarihinde <http://www.hutchinsweb.me.uk/IJT-2004.pdf> adresinden alındı
- Hutchins, W. (2015). Machine Translation: History of Research and Applications. C. Sin-wai (Dü.) içinde, *The Routledge Encyclopedia of Translation Technology* (s. 120-136). Londra & New York: Routledge Taylor & Francis Group.
- Hutchins, W., & Somers, H. (1992). *An Introduction to Machine Translation*. Londra: Academic Press.
- Jermolowicz, R. (2003). On the Project of a Universal Language in the Framework of the XVII Century Philosophy. *Studies in Logic, Grammar and Rhetoric, The Journal of University of Bialystok*, 6(19), 51-61. Ekim 18, 2018 tarihinde <http://logika.uwb.edu.pl/studies/download.php?volid=19&artid=rj&format=PDF> adresinden alındı
- Koehn, P. (2010). *Statistical Machine Translation*. Cambridge: Cambridge University Press.
- Lavie, A., Langley, C., Waibel, A., Pianesi, F., Lazzari, G., Coletti, P., . . . Balducci, F. (2001). Architecture and Design Considerations in NESPOLE!: A Speech Translation System for E-commerce Applications. *Proceedings of Human Language Technologies*, (s. 31-34). USA. doi:10.3115/1072133.1072140
- Papineni, K., Roukos, S., Ward, T., & Zhu, W.-J. (2002). BLEU: A Method for Automatic Evaluation of Machine Translation. *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics* (s. 311-318). Philadelphia: ACL. Kasım 15, 2018 tarihinde [https://www.researchgate.net/publication/2588204\\_BLEU\\_a\\_Method\\_for\\_Automatic\\_Evaluation\\_of\\_Machine\\_Translation](https://www.researchgate.net/publication/2588204_BLEU_a_Method_for_Automatic_Evaluation_of_Machine_Translation) adresinden alındı
- Pietra, S. A., & Pietra, V. J. (1994). Candide: A Statistical Machine Translation System. ARPA (Dü.), *ARPA Human Language Technology Workshop* içinde (s. 457). Plainsboro, New Jersey: Morgan Kaufmann. Eylül 27, 2018 tarihinde <http://aclweb.org/anthology/H94-1100> adresinden alındı
- Qun, L., & Xiaojun, Z. (2015). Machine Translation: General. C. Sin-Wai içinde, *The Routledge Encyclopedia of Translation Technology* (s. 105-119). Londra & New York: Routledge Taylor & Francis Group.
- Rayner, M., Alshawi, H., Bretan, I., Carter, D., Digalakis, V., Gambäck, B., . . . Samuelsson, C. (1993). A Speech to Speech Translation System Built from Standard Components. *Proceedings 1st ARPA Workshop on Human Language Technology* (s. 217-222). Princeton, NJ: Morgan Kaufmann. Kasım 13, 2018 tarihinde <http://aclweb.org/anthology/H93-1042> adresinden alındı
- Romanov, Y., Zaysteva, N., Zelko, V., Romanov, A., Tovmach, Y., Anikina, N., . . . Putin, V. (2003). Artificial Intelligence and Machine Translation in the XXIst Century. *Journal of Quantitative Linguistics*, 10(3), 215-238. Eylül 18, 2018 tarihinde <https://www.tandfonline.com/doi/abs/10.1080/09296170512331383575> adresinden alındı
- Somers, H. (2003). Introduction. H. Somers, & H. Somers (Dü.) içinde, *Computers and Translation* (s. 1-11). Amsterdam / Philadelphia: John Benjamins Publishnig Company.
- Şahin, M. (2015). Çevirmen Adaylarının Gözünden İngilizce Türkçe Bilgisayar Çevirisi ve Bilgisayar Destekli Çeviri: Google Deneyi. *Çeviribilim ve Uygulamalı Dergisi*(21), 43-60.
- Takezawa, T., Morimoto, T., Sagisaka, Y., Campbell, N., Iida, H., Sugaya, F., . . . Yamamoto, S. (1998). A Japanese-to-English speech translation system: ATR-MATRIX. *The 5th International Conference on Spoken Language Processing, Incorporating The 7th Australian International Speech Science and Technology Conference*, (s. 2779-2782). Sydney. Kasım 16, 2018 tarihinde [https://www.researchgate.net/publication/221492004\\_A\\_Japanese-to-English\\_speech\\_translation\\_system\\_ATR-MATRIX/citations](https://www.researchgate.net/publication/221492004_A_Japanese-to-English_speech_translation_system_ATR-MATRIX/citations) adresinden alındı

- Tak-ming , B. W., & Webster, J. J. (2015). Example-Based Machine Translation. In C. Sin-wai, *The Routledge Encyclopedia of Translation Technology* (pp. 137-151). Londra & New York: Routledge Taylor & Francis Group.
- Ulitkin, I. (2011, 10 1). Computer-assisted Translation Tools. *Translation Journal*, 15(1), 1-8. Retrieved October 1, 2018, from <http://www.translationjournal.net/journal/55computers.htm>
- v. Hahn, W., & Amtrup, J. (1996). Speech-to-Speech Translation: The Project Verbmobil. *Proceedings of SPECOM*, (s. 51-56). St.Petersburg. Kasım 11, 2018 tarihinde <https://pdfs.semanticscholar.org/b4c6/c37c9f3d038359b585d23b012d6ead5da13e.pdf> adresinden alındı
- Waibel, A., Lavie, A., & Levin, L. (1997). JANUS: A System for Translation of Conversational Speech. *Kuenstliche Intelligenz (KI)*(97/4). Kasım 12, 2018 tarihinde <https://www.cs.cmu.edu/~alavie/papers/KI.pdf> adresinden alındı