

İngilizce-Türkçe, Türkçe-İngilizce Makine Çevirisinde

Yazılımların Karşılaştırılması

Ahmet Tarcan¹, Ecevit Bekler²

¹ Dicle Üniversitesi, Eğitim Fakültesi, Yabancı Diller Bölümü, 21280, Diyarbakır

² Dicle Üniversitesi, Yabancı Diller Araştırma ve Uygulama Merkezi, 21280, Diyarbakır

tarcan@dicle.edu.tr, ebekler@dicle.edu.tr

Özet: Geçmişten günümüze bir dilden başka bir dile makine çevirisi alanında pek çok çalışma yapılmış ve bu çalışmalar günümüzün en son teknolojisiyle birleşerek sonuç vermeye başlamıştır. Bilgisayar kullanımı yaygınlaştıkça makine çevirisi programlarına olan ihtiyaç da artmakta ve bu alan hızla büyüyen bir sektör haline gelmektedir. Bu çalışmada uluslar arası ve ulusal bazı makine çeviri programlarıyla farklı cümlelerin İngilizce'den Türkçe'ye ve Türkçe'den İngilizce'ye çevirileri yapılmış, yetersiz yönleri ve artıları sunulmuştur.

Anahtar Kelimeler: Makine Çevirisi, Dil, Cümle, Program, Yazılım

Software Comparison in English-Turkish, Turkish-English Machine Translation

Abstract: A lot of works have been done in the field of machine translation from one language into another from the past to the present and these works, with the latest technology of today, have begun to give fruitful results. The need to the machine translation increases as the use of computer becomes common and this field is becoming a rapidly growing sector. In this study, different sentences have been translated from English into Turkish, and from Turkish into English, using some international and national machine translation programmes, and their positive and negative features have been presented.

Key Words: Machine Translation, Language, Sentence, Programme, Software

1. Giriş

Makine çevirisinin tarihi genelde II. Dünya Savaşı'ndan sonra 1950'lerde başlar. 1954'te yapılan 'Georgetown Deneyi' (Georgetown Üniversitesi ve IBM ortaklığıyla gerçekleştirilmiş bu çeviride IBM 701 bilgisayarı kullanılmıştır. Bu sistemin hafızasında sadece 6 gramer kuralı ve 250 kelime vardı.) altmıştan fazla Rusça cümlenin İngilizce'ye tam otomatik çevirisini kapsıyordu. Bu deney büyük bir başarıydı ve makine çevirisi araştırmasına yatırım dönemini başlattı [6].

Yazarlar, üç veya beş yıl içerisinde makine çevirisi sorununun çözülebileceğini iddia edi-

yordu. Ne var ki, gerçek anlamda ilerleme çok daha yavaştı ve on yıllık araştırmanın hayalleri gerçekleştirilmedi başarısız olduğunu ortaya koyan 1966'daki 'ALPAC (Automatic Language Processing Advisory Committee) Raporu'ndan sonra bu alana yapılan yatırım belirgin ölçüde azaltıldı. Bilgisayar gücünün arttığı ve daha az pahalı olduğu 1980'lerin sonlarından başlayarak 'makine çevirisi istatistikli modelleri'ne daha fazla ilgi gösterilmeye başlandı [6].

Bugün itibarıyla, hem Google hem de Altavista'nın BabelFish online metin ve web sayfası çeviri programlarını çalıştıran 'SYSTRAN' (1968'de Dr. Peter Toma tarafından kurulmuş olup en eski makine çeviri

şirketlerinden bir tanesidir. Amerika Savunma Bakanlığı ve Avrupa Komisyonu için çok sayıda çeviri hizmeti vermiştir.) gibi birkaçı online olmak üzere doğal dili çevirmek için pek çok yazılım programları vardır. ‘Tam otomatik yüksek kalite makine çevirisi’ imkanını veren sistem olmamasına rağmen pek çok sistem makul sayılabilecek bir çıktı vermektedir [6].

2. Günümüz Uygulamaları

Var olan sınırlamalarına rağmen makine çevirisi programları günümüzde dünyanın her yerinde değişik organizasyonlar tarafından kullanılmaktadır. Muhtemelen en büyük kurumsal kullanıcı, kurum içi kullanım için belgelerin ilk taslaklarının büyük bölümünün otomatik çevirisini yapmak için ticari amaçlı makine çeviri sistemi olan ‘SYSTRAN’ın üst düzey versiyonunu kullanan ‘Avrupa Komisyonu’dur [6].

Bir Danimarka çeviri ajansı olan ‘Lingtech A/S’, ‘Trados’ ticari CAT araç sistemine dayalı çeviri hafızasıyla birlikte çalışarak ‘PaTrans’ adlı tecilli, kurallara dayalı makine çeviri sistemini kullanarak 1993’ten beri İngilizce’den Danimarkaca’ya patent uygulamalarını çevirmektedir.

İspanya’nın günlük gazetesi ‘Periodico de Catalunya’, bir makine çeviri sistemiyle İspanyolca’dan İngilizce’ye çevirmektedir.

‘Google’, tecilli istatistiksel makine çeviri motoru kullanarak umut verici sonuçların elde edildiğini bildirmiştir. Bu motor halen, yakında daha fazla dil çiftleri ‘SYSTRAN’ motorundan Google motoruna alınmak üzere, Arapça <-> İngilizce ve Çince <-> İngilizce çevirileri için halen Google Çeviri araçlarında kullanılmaktadır. ‘Uwe Muegge’, İngilizce, Almanca ve Fransızca web sayfalarının tam otomatik, yüksek kalitede makine çevirileri yapmak amacıyla Google motoruyla birlikte ‘kontrollü bir dil’ kullanan bir demo websayfası uygulamaya koymuştur.

Son zamanlarda terörizme yoğunlaşarak, Amerika’daki askeri kaynaklar doğal dil mü-

hendisliğine büyük miktarda para yatırmaktadır. ‘In-Q-Tel’ (özel sektör girişimcileri vasıtasıyla yeni teknolojileri teşvik etmek amacıyla Amerikan İstihbarat Topluluğu tarafından büyük ölçüde finansmanı sağlanan bir risk sermayesi fonu), ‘Language Weaver’ gibi şirketler oluşturmuştur.

Şu an Amerika’daki askeri kesim Arapça, Paştu ve Dari gibi dillerin çevirisine ve işlemine ilgi duymaktadır. Darpa’daki (Defense Advanced Research Projects Agency) Bilgi İşleme Teknoloji Ofisi, ‘TIDES’ ve ‘Babylon Çeviri’ gibi programlara ev sahipliği yapmaktadır. Amerika Hava Kuvvetleri bir dil çeviri teknolojisi geliştirmek için 1 milyon dolarlık bir sözleşme yapmıştır [6].

3. Dünyadaki Başlıca Online Çeviri Programları

Günümüzde basit bir cümleden tutun da paragraflar, hatta web sayfası bile çeviren programlar mevcuttur. Bu program dilleri Batı Avrupa dilleri ve dünyadaki başlıca diller arasında dahi iyi çeviri imkanı verirken Türkçe’nin diğer dillerle olan çeviri hizmeti çok gelişmemiştir. Dünya’da pek çok online çeviri programı olmakla birlikte bugün itibarıyla en yaygın olarak kullanılan ve bilinen çeviri programları Systran ve Babylon’dur.

Systran

36 dil çiftinde çeviri imkanı sunmaktadır. Systran’ın Türkçe çeviri hizmeti yoktur.

Ortaklıkları: Ford Motor Company, Cisco Systems, NCR, DaimlerChrysler Corporation, PricewaterhouseCoopers, Dow Corning Corporation ve diğerleri.

Portalları: Google, AOL, Altavista, Apple’s Sherlock Internet Search, CompuServe, Lycos, OracleMobile.com ve diğerleri.

Hizmet Verdiği Organizasyonlar: Avrupa Komisyonu ve Amerikan İstihbarat Topluluğu.

Aşağıda verilen dillerden İngilizce'ye veya Fransızca'ya ve İngilizce'den ve Fransızca'dan bu dillere çeviri sunmaktadır:

R usça'dan İngilizce'ye (1968), Apollo-Soyuz projesi (Amerikan-Sovyet ortak uzay projesi) için İngilizce'den Rusça'ya (1973), Avrupa Komisyonu için İngilizce kaynağı (1975), Arapça, Çince, Danimarkaca, Hollandaca, Fransızca, Almanca, Yunanca, Hintçe, İtalyanca, Japonca, Korece, Norveççe, Sırp-Hırvatça, İspanyolca, İsveççe, Farsça, Polonyaca, Portekizce, Ukraynaca [5].

Babylon

Türkçe dahil olmak üzere 17 dilde metin çevirisi imkanı sunuyor:

İngilizce, Fransızca, Almanca, İspanyolca, İtalyanca, Portekizce, Japonca, İbranice, Çince (Geleneksel), Çince (Basitleştirilmiş), Hollandaca, Rusça, Korece, Türkçe, Arapça, Farsça, Polonyaca ve Ukraynaca [1].

Babylon 1997'de kuruldu. Formula Vision (25%), the investment arm of Reed Elsevier Publishers (28%), Fishman Group's Monitin Publishers, ('Globes'un sahipleri), RAD Data Communications başkanı Zohar Zisapel (5%) ve (14%) diğer hisse ortakları olmak üzere Babylon'un sahipleridir. Babylon 75 kişi çalıştırmaktadır. Dubik Peer 2006'ın başından beri

Yönetim Kurulu Başkanı'dır.

4. Türkiye'deki Bazı Online Çeviri Programları

Türkiye'de Türkçe'den İngilizce'ye, İngilizce'den Türkçe'ye kelime çevirisi yapabilen birçok yazılım programı bulunmasına rağmen cümle ve metin çevirisi yapabilen programlar çok fazla değildir. ProÇeviri, Simka ve Turing cümle ve metin çevirisinde kullanılan başlıca programlardır. Simka çeviri programı İngilizce'den Türkçe'ye çeviri yaparken, Turing 6.01 Türkçe'den İngilizce'ye ve İngilizce'den Türkçe'ye basit cümleleri çevirmekte ve bu program geliştirilme aşamasındadır.

5. Babylon ve Proçeviri

ProÇeviri

Yazılımına 2000 yılının ortalarında başlanan ProÇeviri, İngilizce'den Türkçe'ye 70.000'den fazla kelime çevirisi yapabilmektedir. Programın satışı, ilk sürümü olan 1.2 ile başlamış ve günümüzdeki versiyonu 1.9 ile en son şeklini almıştır [4].

a) İngilizce - Türkçe

Aşağıdaki çeviriler Babylon 6.0.1 ve ProÇeviri 1.9'un deneme sürümlerinden alınmıştır. Seçilen örnek cümleler karşılaştırılarak hangi online çeviri yazılımının daha iyi çeviri yaptığı irdelenmiştir.

Babylon	Proçeviri	Yorum
I love you. = Ben seni severim.	I love you. = Seni severim.	
I'm walking at the moment. = Ben anda yürüyorum.	I'm walking at the moment. = Şu anda yürüyorum.	İngilizce'de non-progressive fiillerden biri olan 'love' fiilini her iki çeviri programı, Türkçe'ye '-yor' ekiyle çeviremiyor.
They go to the cinema on Mondays. = Onlar sinemaya gider.	They go to the cinema on Mondays. = Onlar, Pazartesi'de sinemaya gider.	
I will drink tea. = Ben çay içeceğim.	I will drink tea. = Çayı içeceğim.	ProÇeviri, belirtili nesne olarak çeviriyor 'çay' kelimesini.
I met him. = Ben onla buluştum.	I met him. = Onu karşıladım.	
They don't go to the cinema on Mondays. = Onlar sinemaya gitmez.	They don't go to the cinema on Mondays. = Onlar, Pazartesi'de sinemaya gitmez.	ProÇeviri programı burada zaman zarfını algılıyor.

Did you go there? = Sen oraya gittin mi? Was there a meeting in Ankara? = Ankara'da bir meeting vardı?	Did you go there? = Oraya gittin mi? Was there a meeting in Ankara? = Ankara'da bir toplantı var mıydı?	'Olmak' anlamına gelen 'be' fiiliyle yapılan bu geçmiş zamanlı cümleyi ProÇeviri programı 'var mıydı?' olarak doğru şekilde veriyor.
The man did the exercise. = Adam egzersizi yaptı. The exercise was done. = Egzersiz yapıldı. The flowers are being watered. = Çiçekler.	The man did the exercise. = Adam, alıştırmanı yaptı. The exercise was done. = Alıştırma, yapıldı. The flowers are being watered. = Çiçekler, sulanıyor.	Belirtili nesne kullandığımız bu örnekte karşılıklar farklı kelimelerle veriliyor. Bu, bize çeviri programlarına farklı kelimeler yüklediğini gösterebilir. İkinci örnekteki pasif cümleyi Babylon daha doğru veriyor. Üçüncü örnekte şimdiki sürekli zamanın pasif halini kullandık. Babylon fiili algılayamıyor.
The men in train looked at me. = Men eğit bana baktı. The men, who were in train, looked at me. = Men, oldu eğit bana baktı.	The men in train looked at me. = Trende adamlar, bana baktı. The men, who were in train, looked at me. = Trende olan adamlar, bana baktı.	Yer zarfı kullandığımız ilk cümleyi ProÇeviri biraz daha anlamlı şekilde çeviriyor. Belirtisiz sıfat cümlecığı kullandığımız ikinci örneği ProÇeviri doğru çevirebiliyor.
I will have talked to him. = Ben onla konuşmuş olacağım.	I will have talked to him. = Onla konuşmuş olacağım.	'Future Perfect' zamanlı bu cümleyi ikisi de doğru veriyor.
If it is hot, I will open the windows. = O sıcaksa, ben pencereleri açacağım. I will open the windows if it is hot. = Ben o sıcaksa pencereleri açacağım.	If it is hot, I will open the windows. = Eğer, sıcaksa, pencereleri açacağım. I will open the windows if it is hot. = Pencereleri açacağım, eğer, sıcaksa.	Şart cümlecığıyle başladığımız bu cümledeki şahıs olamayan ve havayı anlatan 'it' zamirini Babylon 'o' diye görüp Türkçe'ye olduğu gibi çevirme hatasına düşmüştür. İkinci örnekte şart cümlecığının yerini değiştirdik.
Don't wait! = Bekleme! Wait! = Bekle! Be quick! = Çabuk! Come here! = Burada gel!	Don't wait! = Bekleme! Wait! = Bekleme! Be quick! = Çabucak ol! Come here! = Buraya gel!	Olumlu bir emir ifadesi olan 'Wait!' i ProÇeviri negatif olarak algılıyor. Babylon, 'Çabul ol!'daki 'olmak fiilini algılayamıyor. Türkçe'de 'burası, buraya' ve 'burada' anlamını veren 'here' yer zarfını Babylon yanlış ekle ifade ediyor, Türkçe'deki fiil – zarf uyumunu algılayamıyor.
We are going to the ball. = Biz topa gidiyoruz. Look at table 3. = Masa 3 bak.	We are going to the ball. = Topa gidiyoruz. Look at table 3. = Masa 3ü'ne bak.	Hem 'balo' hem de 'top' anlamına gelen 'ball' kelimesiyle yaptığımız bu cümlede her iki program sadece 'top' kelimesiyle karşılık veriyor. Yine hem 'masa' hem de 'tablo' anlamında kullanılabilen 'table' kelimesini her iki program da sadece 'masa' diye çevirerek yanlış çeviriyor. Makine çevirisi kelimenin metin içerisinde kazandığı anlamı algılayamıyor. Aynı zamanda 3'e ifadesindeki '-e' eki doğru verilmiyor.

Table 1: İngilizce'den Türkçe'ye makine çevirisinde Babylon 6.01 ve ProÇeviri 1.9 programlarının kıyaslanması

6. Sonuç

Çeviri, aslında biçimlerin bir dönüşüm işlemidir. Her dönüşümde olduğu gibi, çeviri kaynak metin ile erek-metin arasında anlamsal tam bir eşdeğerlik diye tanımlanabilecek olan bir değişmeyen korunmasını gerektirir. Aynı dilde gerçekleştirilen çeviri, daha önceki bir sözce-nin yeniden düzenlenişiyile yetinir. İki farklı dil arasındaki çeviriye, metinler ve diller arasındaki bir çift bağlantılar kuralı girişimi içinde, bir dilden öbürüne geçişi gerçekleştirmeyi dener. Böylece dilbilimin en önemli konusu olan ‘karşıtlık içinde eşdeğerlik’ sorunu en genel biçimiyle ortaya konmuş olur [2].

Bilgisayar kullanımının bu kadar iyi işlemesinin bir sebebi de, dilin gerçek anlamda yazılımın bir şekli olmasıdır. İnsan beyninin bilgisayar donanımı için model alınması gibi, insan dili de bilgisayar yazılımı için model alınmıştır. Artık, bilgisayarlarla çok yaygın ve yoğun kullanımın yaşandığı son on yıldır bunun ne anlama geldiğini tecrübelerle anlayacak bir konumdayız [3].

Elbette bilgisayarlar insan zekasına asla sahip olamazlar. İnsan beynindeki sinir ağları ve beyin hücreleri insanın düşünmesini, anlamlı ifadeler kullanmasını, kısacası çeviri alanında bilgisayarların asla yapamayacaklarını yapar. Makine çevirisi her ne kadar insanoğlu gibi çeviri yapamasa da bugün itibarıyla anlamlı karşılıklar verebilmesiyle belli bir noktaya gelmiştir. Systran bunun en güzel örneğidir. Dünya’da bu alana pek çok yatırımlar yapılmakta, bunun sonucunda da gittikçe daha anlamlı çeviriler yapan programlar ortaya çıkmaktadır.

1945’te 27 ton ağırlığıyla ve bir odayı kaplayan ilk bilgisayar olma özelliği taşıyan ve yirminci yüzyılın en büyük buluşlarından biri olan ENIAC (Electronic Numerical Integrator and Computer) adlı bilgisayarın yapılmasından kısa bir zaman sonra makine çevirisine geçme

çabaları bilişimsel dilbilim dalını ortaya çıkarmıştır ve bu alan günden güne önem kazanmaktadır. Makine çevirisinin ilk denemeleri her ne kadar başarısız olmuş ve bundan vazgeçilmişse de insanoğlunun yaratıcılığı ve azmi sayesinde tekrar bu alana girilmiş ve bunun meyveleri alınmaya başlanmıştır.

Yukarıda iki çeviri programının karşılaştırılması yapılmıştır. ProÇeviri programının uluslar arası bir çeviri programı olan Babylon’a göre İngilizce’den Türkçe’ye daha iyi çeviri yapabildiği görülmektedir. Türkçe’nin sondan eklemeli bir dil olması ve Hint-Avrupa Dil Grubu’ndan farklı özellikler taşıması sebebiyle Türkçe’nin diğer dillere yapılan makine çevirisinin İngilizce’ye veya Fransızca’ya göre çok fazla gelişme göstermediği ortadadır. Türkçe’nin dünyada gittikçe önem kazanan bir dil olması sebebiyle bu konuda bundan sonra hızlı bir ilerleme kaydedileceği düşünülmektedir.

7. Kaynakça

- [1]. Babylon Translation, <http://www.babylon.com>, Babylon Ltd.
- [2]. Kıran, Z., ve Kıran, A., Dilbilime Giriş, SEÇKİN Yayınevi, Ankara, 2001.
- [3]. Lawler J., ve DRY, H. A., Using Computers in Linguistics, Routledge, New York, 2004.
- [4]. ProÇeviri, <http://www.proceviri.com/main.htm>, Mehmet Cem ÖZMEN
- [5]. Systran Language Translation Technologies, <http://www.systransoft.com/index.html>, SYSTRAN
- [6]. Wikipedia The Free Encyclopedia, http://en.wikipedia.org/wiki/Main_Page, Wikimedia Foundation, Inc.