

**T.C.**  
**İSTANBUL SABAHATTİN ZAİM ÜNİVERSİTESİ**  
**LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ**  
**BİLGİSAYAR MÜHENDİSLİĞİ ANA BİLİM DALI**  
**BİLGİSAYAR BİLİMLERİ ve MÜHENDİSLİĞİ BİLİM DALI**

**MAKİNE ÇEVİRİLERİNDE SORUN TEŞKİL EDEN**  
**EŞ YAZIMLI KELİMELEERİ BULUNDURAN İFADELER**  
**İÇİN BİR KURAL TEMELLİ ÇEVİRİ MODELİ**  
**ÇALIŞMASI**

**DOKTORA TEZİ**

**Nevzat ÇAPOĞLU**

**İstanbul**  
**Eylül - 2021**

**T.C.**  
**İSTANBUL SABAHATTİN ZAİM ÜNİVERSİTESİ**  
**LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ**  
**BİLGİSAYAR MÜHENDİSLİĞİ ANA BİLİM DALI**  
**BİLGİSAYAR BİLİMLERİ ve MÜHENDİSLİĞİ BİLİM DALI**

**MAKİNE ÇEVİRİLERİNDE SORUN TEŞKİL EDEN EŞ  
YAZIMLI KELİMELEİ BULUNDURAN İFADELER İÇİN BİR  
KURAL TEMELLİ ÇEVİRİ MODELİ ÇALIŞMASI**

**DOKTORA TEZİ**

**Nevzat ÇAPOĞLU**

**Tez Danışmanı**  
**Dr. Öğr. Üyesi Yahya ŞİRİN**

**İstanbul**  
**Eylül - 2021**

## TEZ ONAYI

Lisansüstü Eğitim Enstitüsü Müdürlüğüne;

Bu çalışma, jürimiz tarafından Bilgisayar Mühendisliği Ana Bilim Dalında, Bilgisayar Bilimleri ve Mühendisliği Bilim Dalında DOKTORA TEZİ olarak kabul edilmiştir.

Danışman: Dr. Öğr. Üyesi Yahya ŞİRİN

Üye: Dr. Öğr. Üyesi Akhtar JAMIL

Üye: Doç. Dr. Ferdi SÖNMEZ

Üye: Dr. Öğr. Üyesi Muhammed DAVUD

Üye: Dr. Öğr. Üyesi Aydın Tarık ZENGİN

Onay

Yukarıdaki imzaların, adı geçen öğretim üyelerine ait olduğunu onaylarım.

Prof. Dr. Metin TOPRAK

Enstitü Müdürü

## **BİLİMSEL ETİK BİLDİRİMİ**

Doktora tezi olarak hazırladığım “**Makine Çevirilerinde Sorun Teşkil Eden Eş Yazımlı Kelimeleri Bulunduran İfadeler İçin Bir Kural Tabanlı Çeviri Modeli Çalışması**” adlı çalışmanın öneri aşamasından sonuçlandığı aşamaya kadar geçen süreçte bilimsel etiğe ve akademik kurallara özenle uyduğumu, tez içindeki tüm bilgileri bilimsel ahlak ve gelenek çerçevesinde elde ettiğimi, tez yazım kurallarına uygun olarak hazırladığımı, bu çalışmamda doğrudan veya dolaylı olarak yaptığım her alıntıya kaynak gösterdiğimi ve yararlandığım eserlerin kaynakçada gösterilenlerden oluştuğunu beyan ederim.

**Nevzat ÇAPOĞLU**

## ÖN SÖZ

Çalışmalarım süresince çözüm odaklı ve pratik yaklaşımlarıyla yardımlarını esirgemeyen tez danışmanım Dr. Öğr. Üyesi Yahya ŞİRİN, Dr. Öğr. Üyesi Akhtar JAMIL ve Doç. Dr. Ferdi SÖNMEZ hocalarıma teşekkürü bir borç bilirim.

**Nevzat ÇAPOĞLU**

**İstanbul-2021**



**ÖZET**

**MAKİNE ÇEVİRİLERİNDE SORUN TEŞKİL EDEN EŞ  
YAZIMLI KELİMELERİ BULUNDURAN İFADELER İÇİN BİR  
KURAL TABANLI ÇEVİRİ MODELİ ÇALIŞMASI**

**Nevzat ÇAPOĞLU**

Doktora, Bilgisayar Bilimleri ve Mühendisliği

Tez danışmanı: Dr. Öğr. Üyesi Yahya ŞİRİN

Eylül, 2021 – 162 + XVI Sayfa

Doğal diller tam olarak matematiksel formüllerle ifade edilemediği ve sosyal kabullere dayanan kelime veya ifadeler barındırdıkları için makine çevirilerinde bazı anlam belirsizliği durumlarıyla, özellikle yazılı metinlerde, karşılaşılabilmektedir. Bu belirsizlik durumuna en uygun örnek eş yazımlı bulunduran ifadelerdir.

Bu tez çalışması eş yazımlı kelime bulunduran ifadelerin kural tabanlı bir çeviri modeli ile doğru şekilde hedef dile çevrilebilmesinin mümkün olup olmadığı sorusunu araştırmaktadır. Bu maksatla Chiang'ın hiyerarşik ifade tabanlı istatistiksel makine çevirisine benzer bir model geliştirilmiş; ancak geliştirilen sistemin eğitiminde Chiang'ın modelinin aksine paralel metinlerden yararlanılmayıp çevirmen tecrübelerine dayanan kural tabanlı yöntem tercih edilmiştir.

Önerilen çeviri modeli başarı ölçümlerinde, sistemin yapısını ve tasarım felsefesini dışarıda tuttuğu için farklı çeviri motorlarının karşılaştırılmasında kullanılan kara kutu testinin uygun olacağı düşünülmüştür. Bu amaçla Google, Bing Yandex ve tez modeli makine çevirilerinin başarı karşılaştırmalarında kullanılmak üzere eş yazımlılar bulunduran cümleler tespit edilmiştir. Bu cümlelerden elde edilen çeviri çıktılarının karşılaştırılmasında çevirmen değerlendirmesinden yararlanılmıştır. Bunun için 80 maddeden oluşan dördümlü Likert ölçeği oluşturulmuş; bu ölçek yardımıyla A ve B çevirmenlerinden akıcılık ve metne sadık olma noktalarında çeviri skorları istenmiştir. Bu sayede her model için başarı sıralaması oluşturulmuştur. Sıralamada Google 90,57, tez modeli 90,26, Bing 85,19 ve Yandex 83,13 çeviri skoru elde etmiştir. Tez modelinin eş yazımlı kelimelerde dikkate değer başarı göstermiş olmasına rağmen Google, Bing ve Yandex çeviri motorlarına kıyasla ek

belirsizliđi durumlarında yeterince etkili olmadığı görülmüştür. Ancak; eklemek gerekir ki ek belirsizliđi, tezin konusu olan bir eş yazımlı hatası olmayıp sözdizimsel hata türlerindedir.

Araştırma sonucunda Google, Bing, Yandex ve tez modeli makine çevirisi başarı oranlarında anlamlı bir fark gözlenmemiş; ancak çeviri hatalarının üç farklı noktada kümeleniđi görülmüştür. Bunlar, (i) eş yazımlı kelimenin yaygın kullanılan anlamının seyrek kullanılan anlamına tercih edildiđi; (ii) çeviri motorları eğitim verilerinin sınırlı olmasından kaynaklanan eş yazımlı belirsizliđi ve (iii) farklı görevlerde kullanılabilen edatlardan kaynaklanan ek belirsizliđi hata durumları şeklinde özetlenebilir.

**Anahtar Kelimeler:** Makine Çevirisi, Doğal Dil İşleme, Kural Tabanlı Makine Çevirisi, Hiyerarşik İfade Tabanlı Makine Çevirisi, Eş yazımlılar

**ABSTRACT**  
**A RULE-BASED TRANSLATION MODEL FOR THE PHRASES**  
**HAVING HOMOGRAPHS THAT CAUSE DIFFICULTIES IN**  
**MACHINE TRANSLATION**

**Nevzat ÇAPOĞLU**

Ph.D., Computer Science and Engineering

Supervisor: Dr. Yahya ŞİRİN

September, 2021 – 162 + XVI Pages

For a natural language cannot be expressed in exact mathematical formulas or can harbor words or phrases that are based on social acceptance, it is very likely to face some ambiguities in machine translations especially in written texts. The most appropriate example would be the homographs.

The purpose of this study is to investigate the question of whether a rule-based translation model can translate the expressions with homonymous ambiguity into the target language. Therefore, a model similar to Chiang's Hierarchical Phrase-based Translation has been developed. However, in the training of the developed system, unlike Chiang's model, the rule-based method relying on human experience has been preferred instead of parallel texts.

To evaluate the proposed model, the black-box view has been chosen for it is a look at the input and output without taking into account the mechanics of the translation engine. In order to determine the translation success of Google, Bing, Yandex and the thesis model, sentences with homonymous ambiguity were picked out. In measuring MT outputs obtained from these sentences, human evaluation has been applied. In the evaluation process, 4-point Likert-type scale with 80 items has been prepared. Two human raters (A and B) had been instructed to assign points in terms of how intelligible and faithful the target translation outputs of the MT models were and thereby ranked the achievement of the systems such as Google with 90,57; thesis model with 90,26; Bing with 85,19 and Yandex with 83,13 points. Being the second of all four engines with homonymous ambiguities, the thesis model might be showing a notable success but when it comes to attachment ambiguities the system

cannot produce sufficient outputs. However; it should be noted that attachment ambiguity is not a homonymous ambiguity, related to the thesis topic, but a syntactic error type.

In this dissertation we did not observe a significant difference in success rates of Google, Bing, Yandex and the thesis model translations. However their translation failures tend to cluster in three certain areas, such as (i) choosing the widely used meaning of a homograph over the rarely used one; (ii) the homonymous ambiguity resulting from the limited training data and (iii) the wrong translation of the sentences with attachment ambiguity caused by the prepositions with multiple grammatical functions.

**Keywords:** Machine Translation, Natural Language Processing, Rule-based Machine Translation, Hierarchical Phrase-based Machine Translation, Homographs

# İÇİNDEKİLER

TEZ ONAYI.....	i
BİLİMSEL ETİK BİLDİRİMİ .....	ii
ÖN SÖZ.....	iii
ÖZET.....	iv
ABSTRACT .....	vi
İÇİNDEKİLER.....	viii
ŞEKİLLER LİSTESİ.....	xiii
SEMBOLLER ve KISALTMALAR .....	xvi

## BİRİNCİ BÖLÜM

GİRİŞ .....	1
1.1. Tezin Amacı .....	2
1.2. Tezin Kapsamı.....	4
1.3. Sınırlılıklar .....	4
1.4. Varsayımlar .....	5

## İKİNCİ BÖLÜM

MAKİNE ÇEVİRİSİ.....	6
2.1. Tarihçe .....	8
2.2. Doğal Dil İşleme.....	9
2.3. Makine Çevirisi Uygulamaları .....	10
2.4. Makine Çevirisinde Kural Tabanlı Yöntemler .....	13
2.4.1. Doğrudan Makine Çevirisi .....	14
2.4.2. Aktarım Makine Çevirisi.....	15
2.4.3. Aradil Makine Çevirisi.....	16
2.4.4. Kural Tabanlı Makine Çevirilerinin avantaj ve dezavantajları .....	18
2.5. Makine Çevirisinde Derlem Tabanlı Yöntemler .....	19
2.5.1. Örnek Tabanlı Makine Çevirisi .....	20
2.5.2. İstatistiksel Makine Çevirisi .....	25
2.5.3. Nöral Makine Çevirisi.....	46
2.5.4. Derlem Tabanlı Makine Çevirilerinin Avantaj ve Dezavantajları... 54	
2.6. Çeviri Değerlendirmesinde Otomatik Yöntemler.....	55

2.6.1. BLEU / NIST.....	56
2.6.2. WER.....	59
2.6.3. METEOR .....	60
2.7. İnsan Değerlendirmesi .....	61
2.8. İngilizce-Türkçe Makine Çevirisinde Karşılaşılan Güçlükler .....	62
2.8.1. Biçimbilim düzeyi .....	63
2.8.2. Sözcükbilim düzeyi.....	64
2.8.3. Sözdizim düzeyi .....	64
2.8.4. İfade düzeyi .....	69

### ÜÇÜNCÜ BÖLÜM

<b>İNGİLİZCE-TÜRKÇE KURAL TABANLI BİR ÇEVİRİ MODÜLÜ .....</b>	<b>71</b>
3.1 Ayırıştırma Ünitesi .....	76
3.1.1. KD dizgisi sorgulama birimi .....	77
3.1.2. HD liste düzenleyici birimi .....	78
3.1.3. İki dilli sözlük.....	80
3.2. Birleştirme Ünitesi.....	81
3.2.1. Çoklu kod sorgulama birimi.....	82
3.2.2. Birleştirme matrisi .....	82
3.2.3. Kod ve dizgi üretim .....	88
3.2.4. Yapım-çekim eki ilave .....	89
3.2.5. Birleştirme ünitesinde HD üretim aşamaları.....	90
3.2.6. Üretim aşamalarında eş yazımlı kelimeler .....	100

### DÖRDÜNCÜ BÖLÜM

<b>UYGULAMA .....</b>	<b>114</b>
4.1. Önerilen KMMÇ Modelinin Test Aşaması ve Kullanılan Yöntem.....	114
4.2. Çeviri Girdileri İçin Eş Yazımlı Kelimeler ve Örnek Cümleler .....	115
4.3. Çeviri Çıktıları ve Çevirmen Değerlendirme Ölçeği.....	118
4.4. Çevirmen Değerlendirme Puanları .....	119
4.5. Eş Yazımlı Kelime Bulunduran Çeviri Hatalarının İncelenmesi .....	121
4.5.1. Yaygın ifadenin az kullanılan ifadeye tercih edilmesi.....	122
4.5.2. Eş yazımlı belirsizliği ( <i>Homonymous ambiguity</i> ) .....	123
4.5.3. Ek belirsizliği ( <i>Attachment ambiguity</i> ) .....	124

## **BEŞİNCİ BÖLÜM**

**SONUÇLAR ve ÖNERİLER..... 127**

**KAYNAKÇA..... 133**

**EKLER..... 148**

Ek: 1 - İsim İfadeleri Çeviri Değerlendirme Ölçeği (Madde 1 – 8) ..... 148

Ek: 2 - İsim İfadeleri Çeviri Değerlendirme Ölçeği (Madde 9 – 16) ..... 149

Ek: 3 - İsim İfadeleri Çeviri Değerlendirme Ölçeği (Madde 17 – 24) ..... 150

Ek: 4 - İsim İfadeleri Çeviri Değerlendirme Ölçeği (Madde 25 – 32) ..... 151

Ek: 5 - İsim İfadeleri Çeviri Değerlendirme Ölçeği (Madde 33 – 40) ..... 152

Ek: 6 - Fiil İfadeleri Çeviri Değerlendirme Ölçeği ( Madde 41 – 48)..... 153

Ek: 7 - Fiil İfadeleri Çeviri Değerlendirme Ölçeği (Madde 49 – 56)..... 154

Ek: 8 - Fiil İfadeleri Çeviri Değerlendirme Ölçeği (Madde 57 – 64)..... 155

Ek: 9 - Fiil İfadeleri Çeviri Değerlendirme Ölçeği (Madde 65 – 72)..... 156

Ek: 10 - Fiil İfadeleri Çeviri Değerlendirme Ölçeği (Madde 73 – 80)..... 157

Ek: 11 - Google, Bing, Yandex ve KMMÇ çıktıları tablosu (Madde 1 – 21)..... 158

Ek: 12 - Google, Bing, Yandex ve KMMÇ çıktıları tablosu (Madde 22 – 40)..... 159

Ek: 13 - Google, Bing, Yandex ve KMMÇ çıktıları tablosu (Madde 41 – 61)..... 160

Ek: 14 - Google, Bing, Yandex ve KMMÇ çıktıları tablosu (Madde 62 – 80)..... 161

**ÖZGEÇMİŞ ..... 162**

## TABLULAR LİSTESİ

Tablo 1.1: Google, Bing ve Yandex yanlış çeviri örnekleri.....	3
Tablo 2.1: ÖTMÇ’lerde tercüme için gerekli çeviri eşlerine örnek bir gösterim.....	22
Tablo 2.2: Farklı yapılar sergileyen İngilizce-Türkçe çeviri eşleri .....	23
Tablo 2.3: Aynı kelime grubunun (nesne) üç ayrı hal eki ile temsili .....	67
Tablo 2.4: Aynı kelimenin üç farklı çeviri örneği .....	68
Tablo 2.5: Tek bir yapının üç farklı çeviri örneği .....	69
Tablo 3.1: Türkçede fiillere göre değişen ismin beş hâli .....	90
Tablo 4.1: Eş yazımlı kelimeler ve isim ifadeleri içinde kullanıldıkları cümleler ...	116
Tablo 4.2: Eş yazımlı kelimeler ve fiil ifadeleri içinde kullanıldıkları cümleler .....	117
Tablo 4.3: Çevirmen derecelendirme ölçütü .....	119
Tablo 4.4: Google, Bing, Yandex ve KMMÇ modeli çevirmen değerlendirme puanları ve her bir modelin çeviri skorlarını gösterir tablo.....	120
Tablo 4.5: Eş yazımlı kelimenin yaygın kullanılan anlamının seyrek kullanılan anlamına tercih edildiği çeviri örnekleri .....	122
Tablo 4.6: Eş yazımlı kelimenin ilişkilendirilecek anlamlarının fazla olmasından (Homonymous ambiguity) kaynaklanan yanlış çeviri örnekleri .....	124
Tablo 4.7: Ek belirsizliklerine (Attachment ambiguity) dayanan yanlış çeviri örnekleri .....	125
Tablo 5.1: Eş yazımlı kelimelerin isim ifadelerinde kullanıldığı cümle çevirilerinin Google, Bing, Yandex, KMMÇ modeli çeviri skorları ve KMMÇ ile diğer modellerin çeviri skoru karşılaştırması.....	128

Tablo 5.2: Eş yazımlı kelimelerin fiil ifadelerinde kullanıldığı cümle çevirilerinin Google, Bing, Yandex, KMMÇ modeli çeviri skorları ve KMMÇ ile diğer modellerin çeviri skoru karşılaştırması .....129

Tablo 5.3: Tüm eş yazımlı kelimelerin isim ve fiil ifadelerinde kullanıldığı cümle çevirilerinin Google, Bing, Yandex, KMMÇ modeli çeviri skorları ve KMMÇ ile diğer modellerin çeviri skoru karşılaştırması .....130



## ŞEKİLLER LİSTESİ

Şekil 2.1: Makine Çevirisi yaklaşımları.....	12
Şekil 2.2: KTMÇ modellerini gösteren Vauquois Üçgeni (Vauquois, 1968) .....	14
Şekil 2.3: Doğrudan Makine Çevirisi modeli.....	15
Şekil 2.4: Aktarım Temelli çeviri modeli (Saini ve Sahula, 2015) .....	16
Şekil 2.5: Aradil yönteminde kaynak ve hedef dillerin gösterimi .....	17
Şekil 2.6: ÖTMÇ'ler için uyarlanmış Vauquois üçgeni .....	21
Şekil 2.7: Gürültülü Kanal yaklaşımının temsili görseli.....	26
Şekil 2.8: İstatistiksel Makine Çevirisi bileşenleri .....	28
Şekil 2.9: Örnek cümle üzerinden trigram parametreleri.....	31
Şekil 2.10: İspanyolca-İngilizce ÇM olasılık puanları .....	32
Şekil 2.11: İspanyolca-İngilizce ÇM ve DM olasılık puanları.....	33
Şekil 2.12: İMÇ'lerde birden fazla kelimeli yapıların hizalanması.....	33
Şekil 2.13: İMÇ'lerde kelime ekleme-çıkarma durumlarında hizalama sorunu .....	34
Şekil 2.14: Kelime tabanlı ve ifade tabanlı çeviri karşılaştırması .....	37
Şekil 2.15: İMÇ'lerde Faktörlü çeviri modeli .....	39
Şekil 2.16: Sözdizim tabanlı İMÇ'lerde işlem aşamaları .....	41
Şekil 2.17: Senkronize bağlamdan bağımsız gramer modeli çeviri örneği.....	45
Şekil 2.18: Kodlayıcı-kod çözücü modeli genel görünümü.....	47
Şekil 2.19: Bir ÖSA yapısı ve döngüsü açılmış haldeki görünümü .....	47

Şekil 2.20: Çift Yönlü Özyineli Sinir Ağları.....	48
Şekil 2.21: Dönüştürücü modeli mimarisi .....	52
Şekil 2.22: Kelimelerin önceki kelimelerle olan ilişkisi gösterimi .....	53
Şekil 2.23: 1-gram kesinlik örneği.....	57
Şekil 2.24: 2-gram kesinlik örneği.....	58
Şekil 2.25: Ünlü uyumunun silsile şeklinde çalışması örneği.....	63
Şekil 2.26: Çeviride öge sıralamalarına göre değişen anlamlar .....	65
Şekil 2.27: Kök kelimenin yalnız soldan nitelenmesiyle özne-nesne sınırının oluştuğu yapı.....	66
Şekil 2.28: Yüklem ortada olmasıyla özne-nesne sınırının oluştuğu yapı.....	66
Şekil 3.1: SynCFG bileşenleri örneği .....	72
Şekil 3.2: SynCFG’lerde üretim aşamaları ağaç gösterimi .....	74
Şekil 3.3: KMMÇ modelinin Vauquois üçgeni üzerinden aşamaları .....	75
Şekil 3.4: KMMÇ modelinin genel görünümü.....	75
Şekil 3.5: Ayrıştırma ünitesi ve bileşenleri .....	77
Şekil 3.6: Birleştirme ünitesi genel görünümü .....	81
Şekil 3.7: Birleştirme matrisi liste elemanları ve irtibatlı oldukları birimler .....	83
Şekil 3.8: Komut karakteri yardımıyla kod ve dizgi üzerinde yapılan değişiklikler ..	85
Şekil 3.9: Karar kodu (?X_?) uygulama örneği.....	86
Şekil 3.10: Karar kodları uygulama örnekleri .....	87
Şekil 3.11: Kontrol kodları ve HD dizgisi üretimi aşaması .....	92
Şekil 3.12: Ayrıştırma ünitesi ve birleştirme ünitesi aktivite diagramı .....	98

Şekil 3.13: KMMÇ’de eş yazımlı bir kelime için doğru anlam seçimi analiz aşamaları .....	102
Şekil 3.14: Tamlanana göre tamlayanın seçildiği analiz örneği.....	103
Şekil 3.15: Üç farklı eş yazımlı için doğru anlam seçimi analiz aşamaları .....	106
Şekil 3.16: “kitap”, “rezerve et” “kaydet” çeviri adaylarının önünde “işbu” artikeli bulunduğu için üç çeviri ihtimalinden ikincisi olan “rezerve et” dizgisinin silinmesi .....	108
Şekil 3.17: “iş”, “çalıştır” “çalış” çeviri adaylarının önünde “benim” sahiplik sıfatı bulunduğu için üç çeviri ihtimalinden ikincisi olan “çalıştır” dizgisinin silinmesi..	109
Şekil 3.18: “kitap” ve “kaydet” çeviri adaylarının önünde “işbu” artikeli bulunduğu için iki çeviri ihtimalinden ikincisi olan “kaydet” dizgisinin silinmesi.....	110
Şekil 3.19: “iş” ve “çalış” çeviri adaylarının önünde “benim” sahiplik sıfatı bulunduğu için iki çeviri ihtimalinden ikincisi olan “çalış” dizgisinin silinmesi.....	111
Şekil 3.20: “eşya”, “reddet” “itiraz et” çeviri adaylarının önünde “işim” cins ismi bulunduğu için üç çeviri ihtimalinden birincisi olan “eşya” dizgisinin silinmesi ....	112
Şekil 3.21: “itiraz et” ve “reddet” çeviri adaylarının ardında yönelme durumu bildiren “kitaba” dizgisi olduğu için (yalnız belirtme hali ile kullanılabilen) “reddet” dizgisinin silinmesi .....	113

## SEMBOLLER ve KISALTMALAR

<b>BLEU:</b>	Bilingual Evaluation Understudy
<b>Çev.:</b>	Çeviren
<b>CFG:</b>	Context free Grammar
<b>ÇM:</b>	Çeviri modeli
<b>DDİ:</b>	Doğal Dil İşleme
<b>DM:</b>	Dil modeli
<b>Haz.:</b>	Hazırlayan
<b>HD:</b>	Hedef dil
<b>İMÇ:</b>	İstatistiksel Makine Çevirisi
<b>KD:</b>	Kaynak dil
<b>KMMC:</b>	Kural Matrisli Makine Çevirisi
<b>KTMÇ:</b>	Kural Tabanlı Makine Çevirisi
<b>MÇ:</b>	Makine Çevirisi
<b>NMÇ:</b>	Nöral Makine Çevirisi
<b>ÖSA:</b>	Özyineli Sınır Ağları
<b>ÖTMÇ:</b>	Örnek Tabanlı Makine Çevirisi
<b>s:</b>	Sayfa
<b>UKVB:</b>	Uzun Kısa Vadeli Bellek
<b>vb:</b>	Ve benzeri
<b>vd:</b>	Ve diğerleri

# BİRİNCİ BÖLÜM

## GİRİŞ

Bir insanı tanımak için kimlerle birlikte olduğuna bakmak yeterlidir. Bir kelimeyi de anlamak için yalnız sözlüğe değil, cümle içinde hangi kelimelerle birlikte olduğuna bakmak gerekir. Bu benzetme için, doğal dil işleme sahasında yer alan tüm makine çevirisi çalışmalarının kelimelere yaklaşımının özeti olduğu söylemek yanlış olmaz.

Her ne kadar bir doğal dil, makine yoluyla bir başka doğal dile eksiksiz ve doğru bir şekilde çevrilemiyor olsa da bu alandaki çalışmalar halen devam etmekte ve son yıllarda yapay sinir ağlarıyla insan beynini taklit eden Nöral Çeviri (*Neural Translation*) çalışmalarının, bazı eksiklerine rağmen, oldukça başarılı sonuçlar verdiği bilinmektedir. Ancak bu çeviri motorlarında dahi uzun mesafe bağımlılıkları veya uzak yapılara sahip diller söz konusu olduğunda doğru kelime karşılıklarının bulunabilmesinde zorluklar yaşanabilmektedir.

Bu tez çalışmasında makine çevirilerinde sorun teşkil eden eş yazımlı kelime bulunduran ifadelerin sözdizimsel ve anlambilimsel kurallara göre tasarlanmış bir çeviri modeli ile doğru şekilde hedef dile çevrilebilmesinin mümkün olup olmadığı sorusu araştırılmıştır. Bu amaçla temel fikir olarak istatistiksel makine çevirilerindeki (2.5.2) hiyerarşik ifade tabanlı istatistiksel makine çevirisine (Chiang, 2007) benzer bir model geliştirilmiştir. Sistemin eğitiminde Chiang'ın modelinin aksine paralel metinlerden yararlanılmayıp çevirmen tecrübelerine dayanan kural tabanlı yöntem tercih edilmiştir. Tasarlanan bu modelde eş yazımlı kelimeler bulunduran cümle çevirileri yapılarak aynı cümlelerin Google, Bing ve Yandex çevirileri ile çevirmen değerlendirme ölçeği (Ek: 1 – Ek: 10) üzerinden karşılaştırılması sağlanmıştır.

Çalışmada öncelikle makine çevirisi başlığı altında kısaca tarihçesi (2.1) ile doğal dil işleme konusuna değinilmiş (2.2) ve makine çevirisi uygulamaları (2.3) ile ilgili yapılmış olan çalışmalara yer verilmiştir. Tez kapsamında önerilen çeviri modeli ve tanımlamalarının temelini teşkil eden konular makine çevirisinde kural tabanlı (2.4) ve derlem tabanlı (2.5) yöntemler şeklinde iki ana başlık altında ele alınmıştır. Çeviri

sistemlerinin kalite ölçümlerinde kullanılan en yaygın otomatik değerlendirme yöntemleri (2.6) ve insan değerlendirmesi yönteminin (2.7) de mercek altına alındığı bölümün son aşamasında İngilizce-Türkçe makine çevirilerinde genel olarak karşılaşılan güçlüklerle (2.8) değinilmiştir.

Çalışmanın üçüncü bölümünde, temel fikir olarak istatistiksel makine çevirilerindeki (2.5.2) hiyerarşik ifade tabanlı istatistiksel makine çevirisine (Chiang, 2007) benzer tasarlanmış bir İngilizce-Türkçe kural tabanlı çeviri modülü önerilmiş ve iki ana üniteden oluşan bu model ve bileşenleri, Ayrıştırma Ünitesi (3.1) ve Birleştirme Ünitesi (3.2) başlıklarıyla ayrıntılı ele alınmıştır.

Dördüncü bölümde ‘uygulama’ başlığı altında, önerilen çeviri modelinin test aşaması ile bu aşamada kullanılmış olan yöntem (4.1) ele alınmış; modelin test edilmesinde ihtiyaç duyulan çeviri girdileri için eş yazımlı kelimeler ile örnek cümleler (4.2) konusuna değinilmiştir. Söz konusu girdiler kullanılarak elde edilen Google, Bing, Yandex ve tez modeli çeviri çıktıları (Ek: 11 – Ek: 14) mercek altına alınmış; çevirmen değerlendirme ölçeği (4.3) verilerine dayanan çevirmen değerlendirme puanları (4.4) üzerinden eş yazımlı kelime bulunduran çeviri hatalarının incelenmesi (4.5) gerçekleştirilmiştir.

Tezin, problem cümlesi ile iki alt problemi üzerinden sonuçlarının değerlendirildiği sonuçlar ve öneriler başlığında ise Google, Bing, Yandex ve tez modeli makine çevirisi çıktılarında ortaya çıkan hataların hangi noktalarda kümelendiğine vurgu yapılırken araştırma ve uygulamaya yönelik değerlendirmelerde bulunulmuştur.

## **1.1. Tezin Amacı**

Bu tez çalışması makine çevirisinde sık karşılaşılan eş yazımlı kelimelerin bulunduğu cümlelerde ortaya çıkan çeviri hatalarının kural tabanlı bir model çalışması yardımıyla engellenebilmesinin mümkün olup olmadığı sorusuna cevap aramaktadır.

Makine çevirilerinde yanlış çeviriler genellikle bu tür farklı anlamları olan kelimelerde görülmektedir. Bu duruma örnek olarak “A crane flew above the construction crane.” cümlesinin Google, Bing ve Yandex çevirileri Tablo 1.1 ile aşağıda sunulmuştur.

**Tablo 1.1: Google, Bing ve Yandex yanlış çeviri örnekleri**

Örnek cümle	Google	Bing	Yandex
A crane flew above the construction crane.	İnşaat vincinin üzerinde bir vinç uçu.	Bir vinç inşaat vinci üzerinde uçu.	İnşaat vincinin üzerinde bir vinç uçu.

Girişi yapılan kaynak dil cümlesinde hem ‘turna’ hem de ‘vinç’ manalarına gelen ‘crane’ kelimesi bulunmaktadır. Cümlenin hedef dildeki karşılığı “İnşaat vincinin üzerinden bir turna uçu.” olması gerekirken ‘...vinç uçu’ şeklinde çevrildiği görülmektedir (Tablo 1.1).

Bu çalışmada genel problem cümlesi olarak “Makine çevirilerinde sorun teşkil eden eş yazımlı kelime bulunduran KD ifadelerinin sözdizimsel ve anlambilimsel kurallara göre tasarlanmış bir çeviri modeli ile doğru şekilde hedef dile çevrilebilmesi mümkün müdür?” seçilmiştir.

Çalışmanın alt problemleri, iki madde halinde şu şekildedir:

- i. İngilizcede isim ifadelerinde (*noun phrase*) geçen eş yazımlıların Türkçeye doğru şekilde çevrilebilmesi mümkün müdür?
- ii. İngilizcede fiil grubu (*verb phrase*) olarak kullanılan ifadelerde geçen eş yazımlıların Türkçeye doğru şekilde çevrilebilmesi mümkün müdür?

## 1.2. Tezin Kapsamı

Bu tez kapsamında Doğal Dil İşleme (DDİ) sahasında yer alan kural tabanlı ve derlem tabanlı çeviri modelleri mercek altına alınmıştır. Çalışmada çekim ekleri, yapım ekleri, uzun mesafe bağımlılıkları için keskin karar mekanizmalarına sahip kural tabanlı yaklaşım kullanan bir çeviri modeli sunulmaktadır. Bu modelde çıktıların tutarlı ve tahmin edilebilir olmasını, çeviri hatalarının kolay tespit edilmesini ve önlem alınması gereken noktaların çabuk belirlenebilmesini sağlayan bir kural matrisi (Birleştirme matrisi) tasarlanmıştır. Dilbilimsel verinin tek bir matriste tutulduğu bu model, tez çalışmasında Kural Matrisli Makine Çevirisi (KMMÇ) modeli olarak adlandırılmaktadır. KMMÇ modeli iki dilli sözlük dışında herhangi bir destek almayıp istatistiksel yöntemlerde olduğu gibi tek veya iki dilli derlemlere ihtiyaç duymamaktadır. Ancak, KMMÇ’de bir aradil görevi yerine getiren kural matrisinin eğitilmesi ve iki dilli sözlük içeriğinin geliştirilmesi noktasında uzman desteğine ihtiyaç duymaktadır.

Modelin değerlendirilmesi sürecinde içinde eş yazımlı kelime bulunduran 80 adet cümleden oluşan değerlendirme ölçeği (Ek: 1 - Ek: 10) geliştirilmiştir. Değerlendirme ölçeğinin oluşturulması aşamasında çevirmen derecelendirme ölçütünden (Tablo 4.3) yararlanılmıştır.

KMMÇ modeli yalnız, makine çevirisinde sorun teşkil eden eş yazımlı kelime ihtiva eden ifade veya cümlelerin kural tabanlı bir modelle doğru çevirilerinin yapılabilmesi amacıyla geliştirilmiştir. Yani model, bütünüyle otomatik çeviri problemlerine odaklanan ve kendisinden önceki çeviri çalışmalarının kapatamadığı açıkları kapatma motivasyonu ile tasarlanmamıştır.

## 1.3. Sınırlılıklar

Doğal diller tam olarak matematiksel formüllerle ifade edilemedikleri ve sosyal kabullere dayanan ifadeler barındırdıkları için biçimsel ve anlamsal tutarsızlıklar sergileyebilmektedir. Bu durum, tez çalışmasında sunulan KMMÇ modeli dâhil tüm dil çalışmaları için bir güçlük olarak ortaya çıkmaktadır.

Makine çevirisinde kalite düzeyi sistem çıktılarının insan çevirisine ne kadar yakın olduğuyla ilişkilidir. McMillan'a (2000) göre araştırmaların değerlendirilmesinde en önemli unsur veri, veri analizi ve sonuçlarının inandırıcı ve güvenilir olmasıdır. Bu nedenle sunulmakta olan çeviri modeli çalışmasında hedef dil çıktılarında kontrol edilmesi gereken 'yeterlilik' ve 'akıcılık' noktalarının ölçülmesinde insan değerlendirmesinden faydalanılmıştır. Çok daha az güvenilir olmasına rağmen makine çevirilerinde otomatik değerlendirme yöntemlerinin yaygın şekilde kullanılıyor olması hızlı ve düşük maliyet tercihlerinden kaynaklanmaktadır. Çalışmada önerilen makine çevirisi modelinin çıktıları Google, Bing ve Yandex çeviri çıktılarıyla karşılaştırılmaktadır. Bu karşılaştırmalarda kullanılan çevirmen derecelendirme ölçütü, 'Çeviri Kalitesinde İnsan Değerlendirmesi' başlıklı bölümde (2.7) sunulmuştur.

Çeviri modülünün her aşamasında liste veri yapılarından yararlanılmaktadır. Bu nedenle programlama dili olarak Python tercih edilmiştir. Çünkü Python liste veri yapısı içerisinde bulunan elemanları aramak, dizgilerde ekleme veya silme gibi işlemler yapmak oldukça kolaydır. Modülde ilaveten ihtiyaç duyulan iki dilli sözlük verileri için SQLite veritabanı ve SQL programlama dilinden yararlanılmıştır.

#### **1.4. Varsayımlar**

Bir doğal dilin başka bir doğal dile doğru şekilde çevirisinin yapılabilmesi için her iki dilin dilbilimsel ve kültürel derinliklerine vâkıf çevirmenlere ihtiyaç vardır. II. Dünya Savaşı'ndan günümüze kadar yapılmış olan tüm makine çevirisi çalışmaları bu gerçeğe işaret etmekte; günümüzün teknik imkânlarında yapılmış olan çalışmalar ancak insan çevirilerine yakın sayılabilecek çıktılar ortaya koyabilmektedir. Örneğin; derin öğrenmeye dayanan nöral çeviri yöntemlerinin oldukça başarılı sonuçlar verdiği bilinmektedir. Ancak bu çeviri motorlarında dahi uzun mesafe bağımlılıkları veya uzak yapılara sahip dillerde doğru kelime karşılıklarının bulunabilmesinde zorluklar yaşanmaktadır.

## İKİNCİ BÖLÜM

### MAKİNE ÇEVİRİSİ

Bu bölümde Makine Çevirisi (MÇ) kısa tarihçesiyle (2.1) ele alınarak Doğal Dil İşlemenin (2.2) hangi konuları içerdiğine değinilmiş ve makine çevirisi uygulamaları (2.3) ile ilgili farklı araştırmacılar tarafından yapılan çalışmalara yer verilmiştir. Araştırmada, literatürde mevcut çeviri yaklaşımları kural tabanlı ve derlem tabanlı yöntemler şeklinde iki ana başlıkta sıralanmıştır. Kural tabanlı yöntemler (2.4); (i) Doğrudan Makine Çevirisi (2.4.1), (ii) Aktarım Makine Çevirisi (2.4.2), (iii) Aradil Makine Çevirisi (2.4.3), Kural Tabanlı Makine Çevirilerinin Avantaj ve dezavantajları (2.4.4); ve derlem tabanlı yöntemler (2.5) ise; (i) Örnek Tabanlı Makine Çevirisi (2.5.1), (ii) İstatistiksel Makine Çevirisi (2.5.2), (iii) Nöral Makine Çevirisi (2.5.3), Derlem Tabanlı Makine Çevirilerinin Avantaj ve dezavantajları (2.5.4) alt başlıklarıyla ele alınmıştır. Yöntemlerin değerlendirilmesinde nitel araştırma sürecine uygun şekilde kitap, dergi, tez ve makalelerden yararlanılmış; ilgili bölümlerin seçilerek tablo ve şekiller yardımıyla görsel hale getirilmesi sağlanmıştır. Literatürdeki bilgilerin toplanması sürecinde, çalışmada önerilen çeviri modeli için arka plan bilgisi oluşturabilecek tanımlama ve örneklendirmelerin seçilmesine dikkat edilmiştir.

Kural tabanlı ve derlem tabanlı yöntemlerin anlatıldığı her iki bölümde de farklı durumlarda avantajları ve dezavantajlarının olduğuna işaret edilmiştir. Ayrıca, sistemlerin çeviri kalitesinin ölçülmesinde faydalanılan en yaygın otomatik değerlendirme yöntemlerinden (2.6) ve insan değerlendirmesi yönteminden (2.7) bahsedilmiştir. Son olarak, İngilizce-Türkçe makine çevirilerinde genel olarak karşılaşılan güçlükler (2.8) biçimbilim, sözcükbilim, sözdizim ve ifade düzeyinde ele alınmıştır.

Terminolojisinde ‘*Machine Translation*’ olarak tanımlanmış ve Türkçeye de ‘Makine Çevirisi’ olarak geçen kavramın, aslında ne denli kabul edilebilir terim olduğu veya gerçekleştirdiği eylemi ne kadar hakkaniyetle temsil ettiği tartışılabilir (Korkmaz, 2019). Halbuki çeviri işlemi yapan, bir makineden ziyade bilgisayardır. Bu nedenle ‘çeviri teknolojisi’ (*translation technology*) veya ‘çeviri yazılımı’ (*translation*

*software*) gibi terimler anlam bakımından daha kullanışlı olacaklardır (Somers, 2003). Ancak; makine çevirisi terimi ile bilgisayarın donanım ve bileşenleri değil, yazılım bileşenleri de akla gelmelidir. ‘makine çevirisi’ ifadesine Türkçe yazılmış akademik metinlerde sık rastlanmakla beraber “bilgisayar çevirisi” ifadesini de görmek mümkündür (Şahin, 2015).

Çeviri kelimesine gelince; Türkçeden yabancı dillere veya yabancı dillerden Türkçeye sözlü veya yazılı bir metnin aktarımını yapan kişiye çevirmen, Türk yazı dillerinden birbirine metin düzeyindeki aktarmalar yapana ise aktarıcı denmektedir. Çeviri veya aktarma yapan kişi kaynak ve hedef dillere hâkim olmanın yanı sıra genel filolojik yeterliğe ve özel bir üretebilme kabiliyetine sahip olmalıdır. Çünkü çevirmenlik ikinci bir üretim sanatı; aktarma işi ise ayrı bir yazarlık işidir. (Musaoğlu, 2003). Musaoğlu’na (A.g.e.) göre bir metnin bir dilden başka bir dile aktarımı, çeviri; metinlerin yazı dilleri arası ve dönemler arası aktarımı ise- çeviri değil- bir aktarma işlemidir. Bu durumda ‘*Makine çevirisi*’ mi yoksa ‘*makine metin aktarımı*’ mı sorusu akla gelse de asıl üzerinde durulması gereken bu ölçekte bir kabiliyete sahip makinenin varlığından söz etmenin günümüzün tüm teknik imkanlarına rağmen mümkün olamadığıdır. Bu nedenle bir ‘çevirmen’ veya bir ‘metin aktarıcısı’ kadar becerikli tercüme yapabilmeyen makine çevirilerinden bahsedilemez.

Makine çevirilerinde bir dilin (kaynak dil) diğer bir dile (hedef dil) otomatik olarak dönüştürülmesi işlemi Doğal Dil İşleme konusunun yarım asrı geçkin süredir ilgilendiği konulardandır. Bunun gerçekleşebilmesi için bilgisayarın her iki dili; o dillerdeki eş anlamlı kelimeleri, kelime öbeklerini ve dilbilgisi kurallarını bilmesi gereklidir (Durgar El-Kahlout ve Oflazer, 2006). Günümüz imkanları dahilinde yeterli sayılacak seviyede bir makine çevirisi için şu üç özelliği olması beklenmektedir:

- Otomatik olması; yani insan müdahalesiz netice alınabilmesi;
- Çeviri kalitesi; yani HD (hedef dil) çıktıları anlaşılabilir ve KD’den (kaynak dil) alakasının olmaması.

- Geniş kapsamlı olması; yani sistemin her türlü konuyu içeren metinler üzerinde işlem görebilmesidir: haber, makale, hikâye, mektup vb.

Bu üç özellik Türkçede ‘Sınırsız Metnin Yüksek Kalitede Otomatik Üretimi’ olarak çevrilebilecek *FAHQT (Fully Automatic-High Quality Output-unrestricted Text)* olarak bilinmektedir (Tantuğ, 2007).

Günümüzde tüm gelişmiş özelliklerine rağmen makine çevirisinin yapabildikleri kalite olarak insan çevirilerinin yanına yaklaşamamakta ve kaynak ve hedef dil metinlerinin iyileştirilmesi için çevirmenler, ön düzenleme (*pre-editing*) ve son düzenleme (*post editing*) yoluyla bu eksikliği kapatma yoluna gitmektedir. Bu yolla elde edilen çalışmalar hem hızlı hem de kaliteli sonuçlar vermektedir.

## **2.1. Tarihçe**

1940'lı yıllara kadar çevirinin otomasyon problemi üzerine çalışılsa da pratikte geliştirmek ve test etme imkânı olmadığından öncü çalışmalar gerçekleştirilememiştir. 1940'lardan 60'lı yılların ortalarına kadar ilk bilgisayarların üretilmesiyle birlikte makine çeviri sistemlerin geliştirilebilmesi mümkün olmuştur. Bu gelişmeyle birlikte alanla ilgili çalışmaların da ilerleyeceğine yönelik beklentiler yükselmiştir. O yıllarda geliştirilen çeviri yaklaşımların bazıları oldukça naif çalışmalardır, ancak ileride keşfedilecek olan temel çeviri yaklaşımlarının önceden kurulmuş hayalleridir (Poibeau, 2017).

1965 ve 1966 yıllarında Amerikan finansman kurumları tarafından yaptırılan ALPAC (*Automatic Language Processing Advisory Committee*) raporu makine çevirisi alanında dramatik bir etki oluşturmuştur. Rapor, özetle o zamana kadar yapılmış tüm araştırmaları kusurlu ve gereksiz olarak nitelendirmiştir. Raporda bilgisayarları ve dilbilimini birleştiren daha temel araştırmalara olan ihtiyaca vurgu yapılmış; (Poibeau, 2017) otomatik sözlük gibi insan çevirisine yardımcı olabilecek araçların geliştirilmesi yönünde tavsiyeler sıralanmıştır (Hutchins, 2015, s. 122).

Rapor özellikle ABD'deki MÇ çalışmaları üzerinde olumsuz etkisi oluşturmuş ve makine çevirisi alanındaki çalışmalar neredeyse durma noktasına gelmiştir. ALPAC, ABD hükümetinin bilgisayar çevirisi çalışmaları için sağladığı fonun bir nevi sonu olmuş; daha da kötüsü MÇ sektöründe genel bir moral kaybına sebep olmuştur (Wilks, 2009). Tüm bu olumsuzluklara rağmen 70'li yıllarda makine çevirisi alanındaki çalışmalara devam edilmiş; Systran projesi AB tarafından satın alınıp uygulamaya konmuştur. Yine Avrupa'da Globalink, Eurotran gibi projeler de devam ettirilmiştir. Ancak istenilen başarı bir türlü elde edilememiştir. Saygın bir bilim insanı olan Janet Pak'ın makine çevirisini imkânsız kılan otuz sayfalık bir raporunu yayınlaması tüm yatırımlara son noktayı koymuştur. Bundan sonra Amerika ve Kanada gibi ülkeler başta olmak üzere yatırımcılar bu rapor doğrultusunda çalışmalarını tamamen rafa kaldırmışlardır (Mihribani, 2006). 1980'lerin sonuna kadar makine çevirisi çalışmaları Avrupa'daki birkaç belli başlı proje dışında pek verimli geçmemiştir.

1990'lı yıllarda hayatımıza 'İnternet' girmiş ve makine çevirisi anlayışında büyük değişiklikler olmuştur. Bilgisayarların hızlanması ve internet verisinin artmasıyla birlikte (Calli, 2016) oldukça geniş iki dilli külliyatlara dayanan istatistiksel çeviri yaklaşımlar ortaya çıkmaya başlamıştır. 2000'li yıllara gelindiğinde ise, web üzerinden otomatik çeviriye duyulan büyük ihtiyaç arafta geçen yaklaşık çeyrek asırlık bir yüzleşme döneminden sonra makine çevirisini tekrar bilgisayarlı dilbilimin en gözde araştırma alanlarından biri haline getirmiştir (Poibeau, 2017).

Günümüzde bu alana en fazla yatırım yapan şirketlerden biri kuşkusuz Google'dır. 2016 yılında hizmete sunduğu Nöral çeviri ile Türkçeyi de sistemine dâhil eden şirket, yüzden farklı dilde makine çevirisi hizmeti sağlamakta ve geleceğe dair bu yöndeki çalışmaları ile büyük umutlar vaat etmektedir (Aslan, 2018).

## **2.2. Doğal Dil İşleme**

Doğal dil, insanın kendisini ifade etmek ve iletişim kurmak için kullandığı bir araçtır. Chomsky (1986) dilin çocuk yaşlarda duyulandan, konuşup anlayabildiği doğal dile dönüşümüne giden süreci insanın genetik yapısıyla alakalı olduğunu söylemektedir.

Makineler için böyle bir yapıdan söz etmek mümkün değildir. Ancak bu bilim dalı doğal dili incelerken her zaman insanın dili edinmesi, anlaması ve konuşmasını şablon almaya gayret etmektedir.

Makine çevirisi uygulamalarını içine alan Doğal Dil İşleme (DDİ) çalışmaları, yapay zeka teknolojilerinden büyük ölçüde faydalanmaktadır. Dilin fonolojik (sesbilimsel), morfolojik (biçimbilimsel), sentaks (sözdizimsel), semantik (anlamsal) özelliklerini inceleyerek çözümleme, yorumlama ve üretme gibi eylemleri gerçekleştirebilecek yazılımlar tasarlamak (Aslan, 2019); yazılı metinlerde yardımcı araçlar geliştirilmesi, metinlerde yazım hatalarının düzeltilmesi, metin bulma ve değiştirme, basılı metni optik olarak okuma ve yanlışlarının düzeltilmesi, metin özeti çıkarma, metinden bilgi çıkarma, metinden herhangi bir bilgiye erişme, metin anlama, bilgisayarda sesle etkileşim, bilgisayarın metni seslendirmesi, konuşmaları metin şeklinde sunma, soru-cevap dizgeleri oluşturma, yabancı bir dili okumada yardımcı araçlar, yabancı dil ile yazma araçları ve doğal diller arası çeviri gibi görevleri amaç edinmiş bir bilim dalıdır (Adalı, 2012).

### **2.3. Makine Çevirisi Uygulamaları**

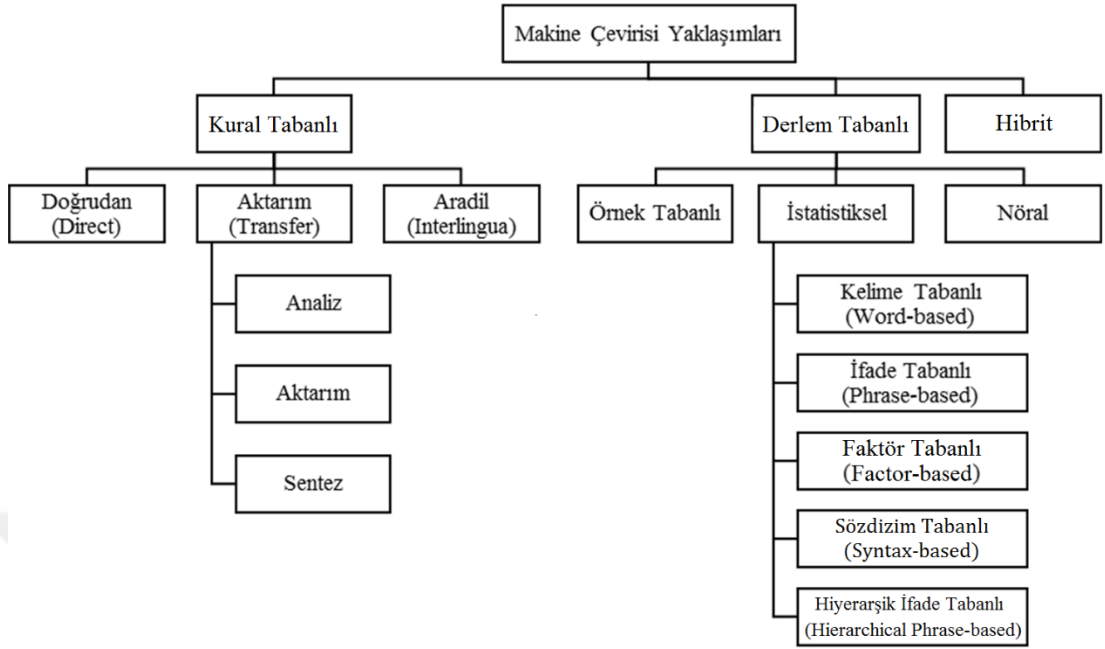
Araştırmanın literatür taraması sürecinde makine çevirisi yöntemlerinin farklı araştırmacılar tarafından farklı şekillerde tanımlandığı görülmüş; alan üzerine yazılmış olan kitap, makale ve tezlerdeki temel tanım ve terminolojilerin farklılıklar arz ettiği tespit edilmiştir.

Aslan (2017) '*Makine Çevirisi*' adlı eserinde bu farklılıklardan bahsederken şu şekilde örneklendirmektedir:

- Tripathi ve Sarkhel (2010) makine çevirilerini (i) Sözlük Tabanlı, (ii) Kural Tabanlı, (iii) Bilgi Tabanlı, (iv) Derlem Tabanlı makine çevirisi olmak üzere dört başlıkta sıralamış;

- Pym (2013) Makine Çevirisi yaklaşımlarını (i) Kural Tabanlı, (ii) Bilgi Tabanlı (*Knowledge-based*), (iii) İlke Tabanlı (*Principle-based*), (iv) Deneysel (*Empirical*), (v) Çevrimiçi İnteraktif, (vi) Hibrit olarak altı farklı kategoriye ayırmış;
- Okpor (2014) (i) Kural tabanlı, (ii) Derlem tabanlı, (iii) Hibrit makine çevirisi olarak üç ayrı başlıkla sunmuş;
- Oladosu vd. (2016) ‘Tek’ ve ‘Hibrit’ ana başlıkları altında ‘Tek’ için (i) Kural tabanlı, (ii) Doğrudan, (iii) Derlem tabanlı ve (iv) Bilgi tabanlı; Hibrit için ise (i) Kelime Tabanlı (*Word-based*), (ii) İfade Tabanlı (*Phrase-based*), (iii) Sözdizimi Tabanlı (*Syntax-based*) ve (iv) Ağaç Tabanlı (*Forest-based*) çeviri yaklaşımı şeklinde sekiz alt başlıkla vermiş;
- Kituku vd. (2016) ise (i) Kural Tabanlı (*Rule-based*), (ii) Karma (Hybrid) makine çevirisi ve (iii) Veri Odaklı (*Data Driven*) makine çevirisi olarak üç ana başlıkta toplamıştır.

Bu alanda yapılan çalışmaların farklılıklar gösteriyor olması bir dezavantaj olarak görülmemelidir. Ancak Silverman’a (2001) göre içerik analizi yapılırken araştırmaların konusu ile alakalı toplanan tüm veriler, temel tanımlar ile açıklanabilir ve karmaşaya mahal vermeyecek kavramlarla izah edilebilir olması gerekir. Bu maksatla ekseriyetle ortak tanımlamalar içeren çalışmalar dikkate alınarak benzer kategorilerde sınıflandırılmış makine çevirisi yaklaşımları ‘Kural Tabanlı’, ‘Derlem Tabanlı’ ve ‘Hibrit’ olmak üzere üç ana başlık altında toplanarak Şekil 2.1 ile görsel hale getirilmiştir.



**Şekil 2.1: Makine Çevirisi yaklaşımları**

Makine Çevirisi (MÇ) sistemleri temelde iki farklı yöntem kullanılarak oluşturulur: Bunlar kural tabanlı ve derlem tabanlı modellerdir. Bilgiye dayalı modeller insan tarafından oluşturulan kuralları kullanarak metni çevirirken; derlem tabanlı modeller iki dildeki çeviri örneklerinden otomatik olarak oluşturulan modelleri kullanmaktadır. Her iki yöntemin de farklı durumlarda avantajları veya dezavantajları ilgili başlıklar ( 2.4.4 ve 2.5.4) altında sıralanmıştır.

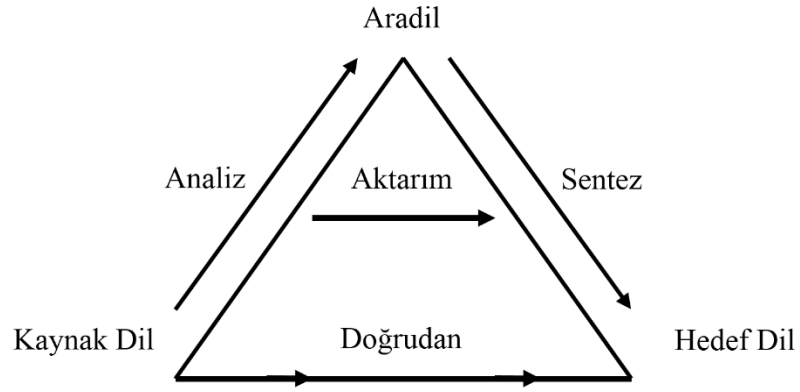
Bu tez çalışmasında kural tabanlı ve derlem tabanlı yöntemler incelenerek kural tabanlı yöntemler (i) Doğrudan Makine Çevirisi, (ii) Aktarım Makine Çevirisi ve (iii) Aradil Makine Çevirisi; derlem tabanlı yöntemler ise (i) Örnek Tabanlı Makine Çevirisi (ii) İstatistiksel Makine Çevirisi ve (iii) Nöral Makine Çevirisi alt başlıklarıyla ele alınmıştır.

#### 2.4. Makine Çevirisinde Kural Tabanlı Yöntemler

Kural Tabanlı Makine Çevirisi, KTMÇ (*Rule-based Machine Translation: RBMT*) 1970'lerde geliştirilmeye başlanan makine çevirisi yöntemlerindedir. KTMÇ uzman kişiler veya dilbilimciler tarafından tasarlanmış (Şahin, 2013: 82) dil kuralları ve sözlüklere dayanan bir sistemdir. Sistemdeki dilbilimsel modüller kaynak metnin **analizini**, iki dil arasında **aktarımı** ve hedef metnin **üretilmesini** veya sentezini sağlamaktadır (Antony, 2013). Bunun için kural tabanlı bir sistemin sözdizimi analizi, anlamsal analiz, sözdizimi oluşturma ve anlamlı yapı oluşturma aşamalarına ihtiyacı vardır: (Okpor, 2014).

KTMÇ sistemlerinin çalışma prensibi verilen girdi cümle yapısının özgün manasını muhafaza ederek istenen çıktı cümlesinin yapısıyla eşleştirilebilmesi üzerinedir (A.g.e.). Böyle bir sistemin oluşturulması kaynak ve hedef diller hakkındaki linguistik bilgi sayesinde gerçekleşir. Bu bilgi temel olarak her iki dilin de anlambilimsel, morfolojik ve sözdizimsel düzenlerini temsil eden gramer yapıları ve sözlüklerden türetilmektedir (Saini ve Sahula, 2015). Dilbilgisi kurallarının makineye öğretildiği bu sistemler literatürde Bilgi Tabanlı Çeviri modeli olarak da anılmaktadır.

KTMÇ Doğrudan (Direct), Aktarım (Transfer) Temelli ve Aradil Makine Çevirileri olarak üç farklı yaklaşımda sınıflandırılmaktadır. Bu mimarileri ele almadan önce yöntemlerini görselleştirmek amacıyla Şekil 2.2 ile gösterilen "Vauquois Üçgeni" aşağıda sunulmuştur.

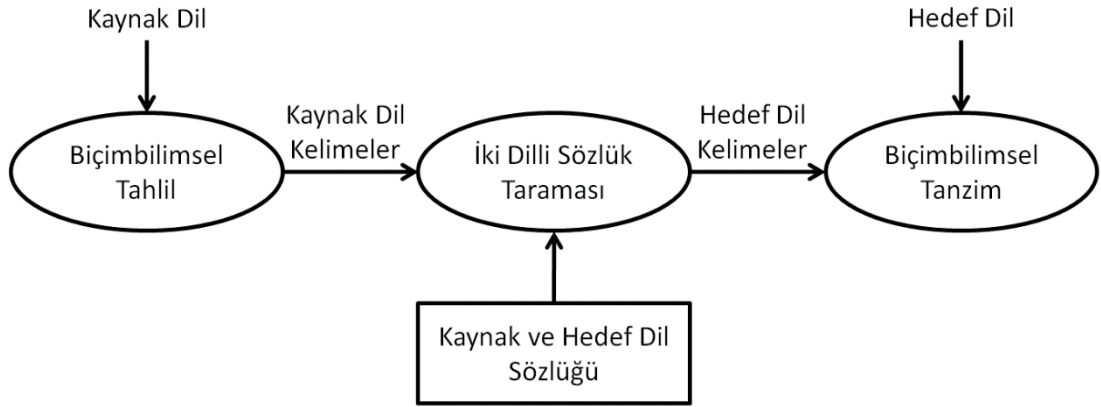


**Şekil 2.2: KTMÇ modellerini gösteren Vauquois Üçgeni (Vauquois, 1968)**

#### **2.4.1. Doğrudan Makine Çevirisi**

Doğrudan makine çevirisinde KD girişinin direkt kelime-kelime çevirisi, ardından bazı sözdizimsel düzenlemelerin de olduğu iki dilli bir sözlük yardımıyla gerçekleştirilir. Sistemde KD girişi yapılarak HD çıktısına ulaşma süreci tek yönlüdür ve tek seferde yalnız bir dil çifti dikkate alınmaktadır. Basit bir çeviri işlemi için sistemin tek ihtiyacı hedef ve kaynak cümle yapılarını temsil eden kurallar ve iki dilde de uygun kelimeleri eşleyebilen bir sözlüktür (Saini ve Sahula, 2015). Bu model, envanterler ve veri yapıları gibi kısa ifadeler barındıran belgelerin çevirilerinde kullanılmıştır (Muegge, 2006).

Doğrudan makine çevirisinde girişi yapılan kaynak dili, biçimsel analiz yoluyla kelimelere ayırıp iki dilli sözlük taraması süresinden geçirdikten sonra hedef dilin sözdizimine göre çıktısının tanzimini gösteren Şekil 2.3 aşağıda sunulmuştur (Saini ve Sahula, 2015):



**Şekil 2.3: Doğrudan Makine Çevirisi modeli**

Kelimeler arasındaki yapı ve ilişkileri dikkate almadığı için özellikle eş anlamlı kelimelerin bulunduğu cümlelerde yanlış çeviriler yapan bu modelin avantajı, benzer dil yapılarına sahip dillerin çevirisindeki hızı ve sistemin kurulum maliyetlerinin düşük olmasıdır (A.g.e.).

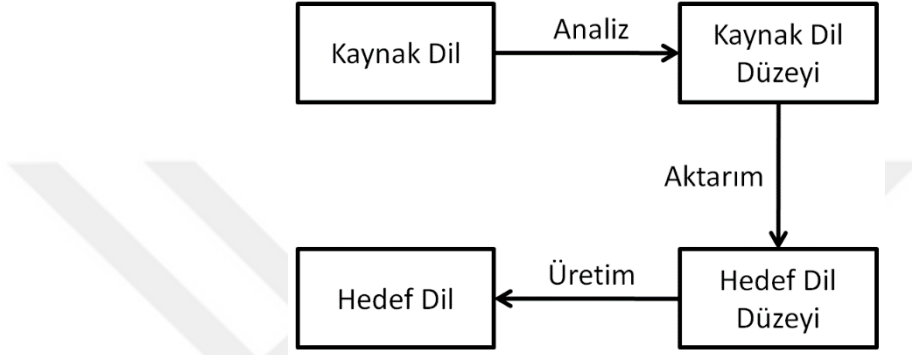
#### 2.4.2. Aktarım Makine Çevirisi

Doğrudan çeviri yaklaşımlarının olumsuzluklarından dolayı geliştirilen ve ikinci nesil olarak nitelendirilen (Aslan, 2017) Aktarım Makine Çevirisi süreci üç adımda gerçekleşmektedir: (i) Kaynak metnin analizi; (ii) sonucun hedef dile uygun yapıda aktarımı (iii) ve hedef dil metninin üretimi.

Sistemin ilk adımı biçim ve kelime çözümlene sürecidir. Bu süreç kelimelerin özne, fiil, sıfat gibi özellikleriyle etiketlenmesidir. Sözcüksel analiz, kelimenin metin içindeki doğru anlamını belirlemek için yapılır. Sistemin ikinci adımı olan aktarım iki bölümden oluşur: sözcüksel ve yapısal. Birincisi doğrudan çeviri modeli gibiyken ikincisi kelimelerin veya cümlelerin yeniden sıralanmasından sorumludur. Dil yapılarının transferi metnin ne kadar yakın diller olup olmadığına bağlı değişiklik göstermektedir. İspanyolca ve Katalanca gibi birbirine benzer yapıdaki diller için sözdizimsel aktarımın kalitesi yeterli görülürken; İspanyolca ve İngilizce gibi yapıca birbirinden daha uzak diller için anlamsal farklılıkları yakalayabilecek daha derin bir aktarım düzeyine ihtiyaç duyulmaktadır. Son olarak üçüncü adımda ise yapısal

aktarım aşamasının çıktıları, uygun biçimlerde HD metni halinde üretilmektedir (Riktors, 2019).

Kaynak metnin analizi; sonucun hedef dile uygun yapıda aktarımı ve hedef dil metninin üretimi süreci Şekil 2.4 ile görselleştirilmiştir.



**Şekil 2.4: Aktarım Temelli çeviri modeli (Saini ve Sahula, 2015)**

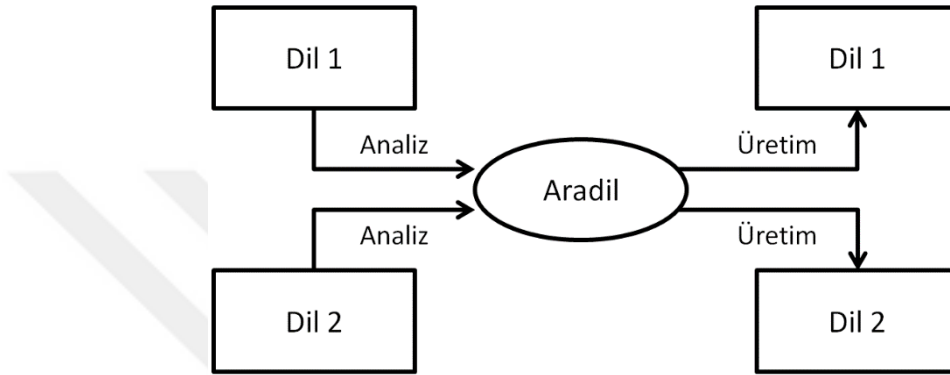
Sistemin iyi yani bir dilden diğerine aktarılan anlam belirsizliklerini yönetebilmesidir. Ancak kaynak metnin hedef metne çevrilmesi sırasında anlamsal bazı kayıplar da olabilmektedir (Dwivedi ve Sukhadeve, 2010) Açık kaynaklı projelerden 'Apertium' aktarım tabanlı çeviri modellerinin en popülerlerindendir (Forcada vd, 2011). Apertium şu anda çevrimiçi olarak yaklaşık 30 farklı dil çifti arasında çeviri yapabilmektedir.

### 2.4.3. Aradil Makine Çevirisi

Tercüman, bir dilden diğerine çeviri yaparken kelimeleri kaynak dilde okur ve zihninde, o dilin yapı ve anlambilim özellikleriyle şekillenmiş bir manalandırma oluşturur. Daha sonra hedef dildeki yapı ve anlambilim bilgilerinden faydalanarak ortaya yeni bir metin çıkarır (Vauquois, 1968). Bu yaklaşımda, KD metni önce bir yardımcı aradile (*Interlingua*) dönüştürülür ve bu aradil aracılığıyla da HD metninin oluşturulması sağlanır (Reddy ve Hanumanthappa, 2012). Yani sistem bir çevirmenin

kaynak dili okuyup anlamlandırması ve ardında hedef dilde ifade etmesine benzetilebilir. Anlamlandırma burada ‘ara lisan’ işlevi görmektedir.

Şekil 2.5 ile Aradil Makine Çevirisi yönteminde farklı dillerin nasıl hedef dillere tercüme edildiği gösterilmektedir.



**Şekil 2.5: Aradil yönteminde kaynak ve hedef dillerin gösterimi**

Kelimeleri anlamlı cümleler haline getirebilmek için kapsamlı bir bilgi kaynağına (kelimeler, gramer kuralları, alana özgü bilgi) ihtiyaç duyulduğundan Aradil yöntemleri genellikle bilgi tabanlı olarak da adlandırılmaktadır.

Aradil yönteminin avantajı anlam temelli bir temsil kabiliyetine sahip olduğu için ‘bilgi verme’ gibi çevrimiçi işlemlerde kullanılabilir olmasıdır. Ancak sistemin zaman verimliliği, uygulanan modelin karmaşıklığıyla doğru orantılı düşüktür. İşin en önemli ve zor kısmı ise cümlelerin anlamını koruyacak soyut bir evrensel dilin (aradil) temsilinin ortaya çıkarılabilmesidir (Saini ve Sahula, 2015). Başarılı ticari aradil MÇ uygulamalarından biri KANT projesidir. Basitleştirilmiş (zamir, bağlaç vb. içermeyen) teknik İngilizce ile yazılmış belgelerin çevirilerinde kullanılmıştır (Nyberg ve Mitamura, 1992).

Genelde Kural tabanlı modeller insan emeğine oldukça fazla ihtiyaç duyan sistemlerdir. Modelde, kaynak metni dilden bağımsız bir anlam temsiline yani aradile dönüştürebilmek için morfoloji ve semantik bilgisine sahip kişilere ihtiyaç

duyulmaktadır. Hatta bu kişilerin hayat tecrübesi dahi bu konuda önemli rol oynamaktadır (Skadins, 2011). Ancak aranan özelliklere sahip uzman kişiler hem kolay bulunamamakta hem de dilin uzmanı da olsa, bir kişinin okuduğu bir metnin kaynak dildeki mana temsilini algılama şekli kişiden kişiye değişebilmektedir.

Kural Tabanlı makine çevirisi ile yapılan tercüme kaynak ve hedef dillerin yapısal benzerlikleri ile doğru orantılı olarak başarılı sonuçlar vermektedir. Bu tür sistemlerde aynı veya benzer aileden gelen doğal diller arası yapılan çeviriler daha gerçekçi sonuçlar vermekte; farklı dil ailelerine bağlı, yapısal özelliklerinde ortaklık göstermeyen diller arasında yapılan çevirilerde olumsuz sonuçlar görülmektedir (Balkul, 2015).

#### **2.4.4. Kural Tabanlı Makine Çevirilerinin avantaj ve dezavantajları**

Doğal dil davranışlarını yakalayabilen bir çeviri sistemi oluşturabilmek için insan müdahalesine ihtiyaç duyan kural tabanlı yöntemlerin üstün tarafları maddeler halinde şu şekilde sıralanmıştır:

- Sistemin eğitilmesi gibi bir durum bulunmadığından büyük paralel derlemlere ve istatistiksel bir model geliştirmeye ihtiyaç yoktur.
- İyi bir sözlük ve kural sistemi oluşturulduktan sonra kaliteli çeviri elde edilebilir. Bunun için sistemin kontrollü bir şekilde dil çalışmalarıyla desteklenmesi yeterlidir.
- Kontrollü dil çalışmaları yapıldığında ihtisas alanlarıyla ilgili cümlelerde yüksek başarı elde edilebilir. Yani çevrilmesi mümkün olmayan mesleki terminolojide, jargonlarda vb. ifadelerde sisteme yapılacak olan dış yardım onu başarı olarak istatistiksel sistemlerin önüne geçirebilir. Bu da belli kullanım alanları için büyük avantaj anlamına gelmektedir. Örneğin; veritabanı çevirisi, kullanma kılavuzu çevirisi vb.
- Çıktılar her zaman tutarlı ve tahmin edilebilirdir. Bu da önlem alınması gereken noktaların belirlenmesi işini kolaylaştırır.

- Çeviri hatalarına neden olan durumların tespitinin kolay olduğu sistemlerdir.
- Çekim ekleri, yapım ekleri, uzun mesafe bağımlılıkları gibi durumlarda keskin karar mekanizmalarına sahiptir. Yapı olarak uzak olan dillerde başarılı oldukları gözlenmiştir.
- Yüksek maliyetli donanım ihtiyacı duyulmadan kaliteli çıktılar elde edilebilir.

Kural tabanlı çeviri modellerin eksik yanları ise maddeler halinde şu şekilde özetlenebilir:

- Oldukça emek yoğunudur ve insan emeğiyle işlenmiş kurallara bağlı sistemler oldukları için eğitilmelerinde yüksek alan bilgisine ihtiyaç vardır.
- İnsan emeğiyle işlenmiş kurallar belli bir dile özgüdür ve benzer olmayan dil çiftlerinde analiz ve aktarım modüllerinin tekrar kullanım imkanı yoktur. Yani farklı diller için yeniden insan emeğine ihtiyaç duyar.
- Sözcükbilimsel ve sözdizimsel araçlar ile çift dilli sözlüklere olan ihtiyaçları nedeniyle kaynak eksikliği olan bazı dillerde kolay uygulanamaz.
- Kültüre ait bazı deyim veya kullanımları ayırt etmeleri zordur.

## **2.5. Makine Çevirisinde Derlem Tabanlı Yöntemler**

Derlem tabanlı çeviriler uzun yıllar çeviri dünyasında en fazla kullanılan metodoloji olarak bilinmektedir (Saini ve Sahula, 2015). Veri güdümlü makine çevirisi olarak da bilinen sistem ana kaynak olarak büyük miktarda iki dilli paralel metin külliyatını (derlem) kullanır (Rikters, 2019). Bu derlem önceden hazırlanmış iki dilde tercüme veya çeviri örnek cümlelerini barındırmaktadır. Bu veriler sayesinde sistem eğitilmekte ve oluşturulan çeviri modelleriyle tercüme işlemleri gerçekleştirilmektedir. (Alsohybe, 2017). Sistem, her iki dilin de istatistiksel çözümlemesini yaparak çeviri görevini yerine getirebilmesi için büyük miktarda veri kümesine yani yeterli büyüklükte bir derleme ihtiyaç duymaktadır (Saini ve Sahula, 2015).

Genellikle aynı sistem kurulumuyla, yalnızca eğitim veri kümesi değiştirilerek birden çok dil çifti için kullanılabilmesi mümkündür. Böylelikle dilbilimsel analiz ve çeşitli kural bileşenleri için insan emeğine duyulan ihtiyaç kayda değer derecede sınırlandırılabilir. Derlem tabanlı çeviri sistemlerinde, yaygın olarak kullanılan diller için gerekli külliyat yeterli miktarlarda bulunabilirken daha az kullanılan diller için temini yetersiz boyutlarda gerçekleşmekte ya da derlem hiç bulunmamaktadır (Rikters M . (2019).

Bu bölümde derlem tabanlı yaklaşımlar (i) Örnek Tabanlı Çeviri, (ii) İstatistiksel Makine Çevirisi ve (iii) Nöral Makine Çevirisi şeklinde üç başlık altında sıralanmış; İstatistiksel makine çevirisi konusu, Kelime tabanlı, İfade tabanlı, Faktörlü, Sözdizim tabanlı, Hiyerarşik İfade tabanlı çeviri olmak üzere beş farklı yöntemiyle ele alınmıştır.

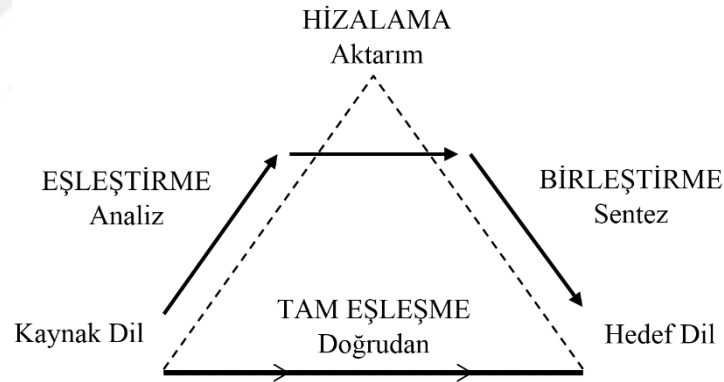
### **2.5.1. Örnek Tabanlı Makine Çevirisi**

Örnek Tabanlı Makine Çevirisi (ÖTMÇ) ilk kez 1980'li yıllarda Japonya'da Makoto Nagao (1984) tarafından tanıtılmıştır. Nagao, o yıllarda hala yaygın olan kural tabanlı sistemlerin zaman içinde giderek daha karmaşık hale gelmekte olduğunu ve işin daha da içinden çıkılmaz hale geleceğini fark etmiştir. Çünkü KTMÇ sistemlerinde genellikle çevrilecek cümlenin tam bir analizine ihtiyaç vardır. Bu da sistemi oldukça zayıf kılmaktadır. Girdisi yapılan bir cümlenin sadece bir kısmı bile analiz edilemezse tüm cümlenin çevirisi imkânsız hale gelmektedir. Halbuki Nagao, tercüme işi ile uğraşan insanların tam ve tutarlı cümleler oluşturmak için cümle parçalarını tercüme ettiklerini ve daha sonra bunları yeniden birleştirdiklerini gözlemlemiş ve çevirmenlerin tercüme edecekleri metinlerin tam bir ön analizini yapmadığını farketmişti. Bu nedenle Nagao, diller arası yeni sözlük, yeni analiz veya farklı transfer yolları keşfetmek yerine doğrudan iki dilli derlemlerde bulunan çeviri parçalarının kullanılmasının daha faydalı olacağını öne sürmüştür (Poibeau, 2017).

ÖTMÇ ile kaynak dildeki cümle kümesinin noktadan noktaya eşleştirilmesi ile hedef dilde karşılığı olan cümlelere dönüştürülmesi hedeflenmektedir (Ballabh ve Jaiswal, 2015) İşlem eşleştirme (*matching*), hizalama (*alignment*) ve birleştirme

(*recombination*) olmak üzere üç vetiredir: (i) **Eşleştirmede** çevrilecek cümlenin derlemde KD için mevcut olan parçaları aranır ve ilgili tüm parçalar toplanarak saklanırken (ii) **hızalamada**, çeviri için kullanılmakta olan eşleştirilmiş tercüme yardımcılarıyla, hedef dildeki çeviri eşdeğeri parçalar bir sırada toplanır ve (iii) son olarak **birleştirme** ise bu çeviri parçalarını hedef dilde doğru bir cümle meydana getirmek üzere bir araya getirir (Poibeau, 2017).

ÖTMÇ modeli yapı olarak üç aşamalı (analiz, aktarım, üretim) olan Aktarım Makine Çevirisi modeliyle (Şekil 2.4) benzerlikler sergilemektedir. Şekil 2.6 ile modelin işlem aşamaları, aktarım tabanlı MÇ sınıflandırmalarıyla birlikte Vauquois üçgeni üzerinde birleştirilerek (Somers, 1999) daha kolay anlaşılır hale getirilmiştir. ÖTMÇ modelinin aşamaları Aktarım yöntemindeki sınıflandırmaların üzerinde büyük harflerle gösterilmiştir.



**Şekil 2.6: ÖTMÇ'ler için uyarlanmış Vauquois üçgeni**

Üç özelliği ile KTMÇ'nin aktarım modeline benzeyen ÖTMÇ'yi Şekil 2.6 üzerinden incelendiğinde (i) 'analiz' aşamasının, arama ve uyanını bulma işlevi gören 'eşleştirme' ile (ii) 'aktarım' oluşumunun, bulup örneğini çıkarma görevini üstlenen 'hızalama' ile (iii) ve 'sentez' sürecinin de metin üretiminin gerçekleştiği 'birleştirme' aşamalarıyla yer değiştirdiği görülmektedir.

Yalnız; Şekil 2.6 içinde ‘Tam eşleşme’ olarak gösterilmiş olan Örnek tabanlı MÇ aşaması, Kural Tabanlı MÇ’deki ‘doğrudan çeviri’ işleminin tam karşılığı değildir. Çünkü ‘tam eşleşme’, mükemmel uyum göstermiş olan cümlelerin adaptasyona uğramadan direkt çevrilme durumları için geçerlidir (Groves, 2007).

Modelin daha net anlatılabilmesi için “*Learning German is not difficult.*” cümlesinin girdi olarak verildiği bir ÖTMÇ sistemi düşünülebilir. Böyle bir girdinin verilmesi durumda sistem benzer kelimeleri barındıran tüm çeviri çiftlerini derlem içinde arayacak ve Tablo 2.1 ile gösterilen ‘*is not difficult*’ ve ‘*learning German*’ parçalarının içinde bulunduğu tüm örnek cümleleri sıralayacaktır.

**Tablo 2.1: ÖTMÇ’lerde tercüme için gerekli çeviri eşlerine örnek bir gösterim**

Örn. 1	This process <i>is not difficult</i> .	Bu işlem <i>zor değildir</i> .
Örn. 2	Speaking German <i>is not difficult</i> .	Almanca konuşmak <i>zor değildir</i> .
Örn. 3	<i>Learning German</i> is easy.	<i>Almanca öğrenmek</i> kolaydır.
Örn. 4	<i>Learning German</i> was fun.	<i>Almanca öğrenmek</i> eğlenceliydi.

Bu örneklerin (Tablo 2.1) hem kaynak hem de hedef dilde tekrarı görülen parçaları ayırt edilmektedir. Bu ayrıştırma neticesinde Örnek 1 ve Örnek 2’de ‘*is not difficult*’ parçasının karşı çevirisindeki tekrarının ‘zor değildir’ olduğu; aynı şekilde Örnek 3 ve Örnek 4’teki ‘*learning German*’ın tekrarının ise ‘Almanca öğrenmek’ olduğu tespit edilecektir. Böylece Örnek 1 ve Örnek 2 yardımıyla ‘X zor değildir’ çevirisi; Örnek 3 ve Örnek 4 yardımıyla da ‘Almanca öğrenmek Y’ çevirisi ortaya çıkmış olacaktır. Bu aşamadan sonra geriye sadece X ve Y’nin birleştirilmesi ve ‘Almanca öğrenmek zor değildir’ metninin üretilmesi kalacaktır.

Ekleme gerekir ki Tablo 2.1 ile verilmiş olan derlem çeviri örnekleri, eşleştirmenin dört çeviri çifti yardımıyla gerçekleştiği basitleştirilmiş bir örnektir. Gerçek uygulamalarda bu derece uyumlu ardışık kelime bütünlerine denk gelmek seyrek görülen bir durumdur (Poibeau, 2017).

Özellikle Türkçe gibi Ural-Altay ailesine mensup diller söz konusu olduğunda çevirilerde kaynak ve hedef dil yapıları arasında büyük farklılıklar ortaya çıkmaktadır. Üç örnek üzerinden Tablo 2.2 ile sunulmuş olan çeviri eşlerinde, İngilizce-Türkçe çevirilerindeki bu farklılıklara işaret edilmektedir.

**Tablo 2.2: Farklı yapılar sergileyen İngilizce-Türkçe çeviri eşleri**

Örnek	Kaynak Dil	Hedef Dil
1	I'm selling the car.	Arabayı satıyorum.
2	I'm selling the car <b>the dog likes</b> .	<b>Köpeğin sevdiği</b> arabayı satıyorum.
3	I'm selling the car <b>the dog I bought likes</b> .	<b>Satın aldığım köpeğin sevdiği</b> arabayı satıyorum.

Tablo 2.2 ile sunulan örneklerde eşleşen yapılar düz, italik ve kalın harf yöntemiyle ayırt edilir hale getirilmiştir. Örnek 2'de '**the dog likes**' cümle yapısı, sıfat fiil olan '**Köpeğin sevdiği**' ifadesine denk gelmekte ve bu yapılar tamamen farklı yerlerde konumlanmaktadır. Örnek 3'te ise öznesi '**the dog I bought**' olan '**the dog I bought likes**' cümlesi, yine bir sıfat fiil olan '**Satın aldığım köpeğin sevdiği**' ifadesiyle eşleşmiştir. Bu iç içe geçmiş ve tarif edilmesi zor durum, aslında dilin kendine has hiyerarşik yapısının sıralı bir şekilde ifade edilmeye çalışılmasından kaynaklanmaktadır.

ÖTMC'nin temel unsurları, cümle kesitleri; temel çalışma tekniği ise giriş dizgilerini veritabanındaki KD dizgileri ile eşleştirmesi; hedef dile karşılık gelen dizgilerin çıkarılması ve makul HD cümleleri olarak yeniden birleştirilmesidir. Ancak sistemde oldukça fazla sayı ve türde tekniklere rastlanmaktadır. Tekniklerin çoğu KTMÇ sistemindeki yöntemler, İMÇ'de bulunan metotlar veya ÇH'de (Çeviri Hafızası)

kullanılan yöntemler gibi; başka modellerden dönüştürülerek elde edilmişlerdir. Yani ÖTMC'nin ne olup olmadığı ile alakalı bir uzlaşmanın olduğu tam da söylenemez (Hutchins, 2005).

Bu yaklaşımın sınırları da ortadadır: Yeterli derecede çeviri parçaları örnekler kümesinde olmadığına sistemin hata vereceği ya da motamot çeviri yapacağı bilinir. Yaklaşım esas itibariyle manuel transfer kurallarının geliştirilmesinin zor olduğu Japonca-İngilizce gibi genetik yapısı uzak diller için geliştirilmiştir. Sistem aynı zamanda sınırlı kelimeleri olan, sıklıkla jargon ve terminoloji içeren belirli bir alt dil kullanan bilgisayar dokümantasyonu gibi çalışma alanları için de kullanılmıştır (Gough ve Way, 2004). Bu tür belirli yapı ve terminolojiler içeren çerçevesi daraltılmış metinlerde, ÖTMC makul denecek yeterlikte metinler ortaya koymaktadır. Sonuç olarak bu çeviri yaklaşımı belli alanlarda ilginç olmasının yanında tek başına kullanılmasının pratikte yeterince verimli olmadığı düşüncesi hâkimdir. Bu nedenle ÖTMC'lerin daha karmaşık sistemlerde modül olarak kullanıldığı; özellikle geniş bir derleme sahip istatistiksel çözümleme ile birlikte kullanıldığında çok olumlu sonuçlar verdiği görülmektedir. Bunda, istatistiksel yaklaşımların çok iyi çağrıştırma yapması ve örnek tabanlı paradigmanın kesinliği büyük rol oynamaktadır (Poibeau, 2017).

ÖTMC'lerin en önemli güçlerinden biri veritabanına daha fazla örnekler eklenerek kapasitesinin artırılabilmesidir. Ama bu her zaman işe yaramaz. Performansı artırmak için eklenmiş veriler, seyrek kullanılan veya yanlış yönlendiren; dolayısıyla da veritabanı kalitesini düşüren örnekler olabilir. Bu durum veritabanının potansiyel tekrar kullanılabilirliğini sınırlayabilir (Somers, 1999). Diğer yandan, bazı diller için yetersiz veri problemi olan derlemlerden de bahsetmek gerekir ki bu konuda yapılan iyileştirme çalışmalarına Vandeghinste vd'nin (2005) METIS projesi örnek gösterilebilir (Hutchins, 2005). METIS sistemi hem KD hem de HD için minimum kaynak ve araçlardan yararlanan, yani herhangi bir paralel derlemeden değil, tek bir HD derleminden yararlanmaktadır (Vandeghinste, 2005).

## 2.5.2. İstatistiksel Makine Çevirisi

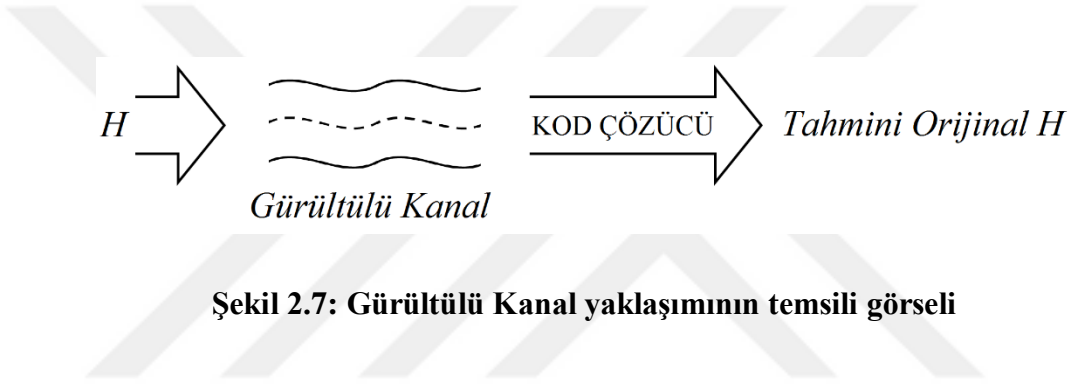
İstatistiksel Makine Çevirisinin (İMÇ) fikir babası Warren Weaver (1894-1978) adlı Amerikalı bir matematikçidir. Weaver (1955) İkinci Dünya Savaşı yıllarındaki kriptografi ve bilgi teorisi alanlarında geniş yer tutan istatistiksel yöntemlerin, metinleri bir dilden diğerine otomatik olarak çevrilmesi işlemi için de uygulanabileceğini öne sürmüştür. Weaver'ın düşüncesi insan dilinde yazılmış her metnin diğer kodlar gibi kırılabilen kodla yazılmış olarak görülebileceği varsayımına dayanıyordu (Britannica, 2020). Ancak o zamanki bilgisayar teknolojisinin yetersizliği, dijital okunur belgelerin olmaması ve bazı karşı görüşler nedeniyle bu erken düşünce hayata geçirilememiş ve çalışmalar transfer ve aradil tabanlı sistemler gibi olasılık teorisinden çok dilbilim üzerine kurulmuş yaklaşımlar lehine hızla terk edilmiştir (Groves, 2007).

MÇ'de istatistiksel yaklaşımın yeniden ortaya çıkışı, ilk istatistiksel yöntem teşebbüslerinden kırk yıl sonrasına denk gelmektedir. İstatistiksel Makine Çevirisi (İMÇ) ana hatlarıyla ilk olarak konuşma tanıma yazılımları üzerinde çalışan araştırmacılar tarafından ortaya konmuştur (Brown vd, 1988, 1990). 80'li yılların sonlarında Yorktown Heights, New York'ta bulunan bir IBM araştırma ekibi, gerçekte konuşma transkripsiyonu için geliştirilen tekniklere dayalı bir makine çevirisi sistemi geliştirmeye karar verdi. 'Konuşma transkripsiyonu' bir ses dizisinden yazılı bir metin üretimi işini ifade etmektedir. Tercüme süreci de buna benzer bir durumdur. Tek fark girişi yapılan verinin bir ses dizisi değil, bir kelime dizisi olmasıdır (Poibeau, 2017).

90'lı yılların başlarında bilgisayarların doğrudan erişebildiği mevcut elektronik metinlerin sayısında büyük artışlar olmuştur. Bu metinlerin bazıları birbirinin çevirisi olan cümlelerdir ve bu çeviriler paragraf veya cümle düzeyinde 'hizalanabilmekte' veya eşleştirilebilmektedir. Bu hizalanmış metinler uzun yıllar çevirmenler için paha biçilmez bir kaynak olmuştur. Kısa bir süre sonra bu yeni kaynaktan bilgisayarlar da yararlanmaya başlamış ve bu gelişme makine çevirisi yaklaşımlarında devrim niteliğinde çalışmaların başlangıcı olmuştur (Poibeau, 2017). İki dilli derlemlerin kullanıma sunulmasıyla birlikte makine çevirisinde yoğun emek gerektiren dilbilime

dayalı kurallı sistemlerin yerini otomatik, hızlı ve kolay eklenilebilir istatistiksel yöntemler almaya başlamıştır (Durgar, 2009).

Geliştirilen paralel derlemlerden yararlanarak, Brown vd. (1993) makine çevirisi için IBM Modelleri adında beş istatistiksel model tanımlamıştır. Bu modeller kelime tabanlı İMÇ prensiplerinin temeli olarak bilinmektedir ve her biri bir öncekinin eksik yanlarını giderebilme motivasyonu ile tasarlanmıştır. Shannon'ın (1948) gürültülü kanal modelinden (Şekil 2.7) yararlanan bu İMÇ'ler oldukça yaygındır (Jurafsky ve Martin, 2000).



Şekil 2.7: Gürültülü Kanal yaklaşımının temsili görseli

Bu yaklaşımda kaynak cümle K, hedef dildeki H cümlesinin gürültülü kanaldan geçtikten sonra bozulmuş hali gibi kabul edilir. Gürültülü Kanal yapılarında kod çözücünün görevi ise bozulmuş olan K'yi ilk hali olan hedef dil cümlesine geri getirebilmektir. Bayes kuramına dayanan bu modeli (Şekil 2.7) temsil etmek için sunulan aşağıdaki denklemde (2.1) hedef dil H, kaynak dil ise K ile gösterilmiştir.

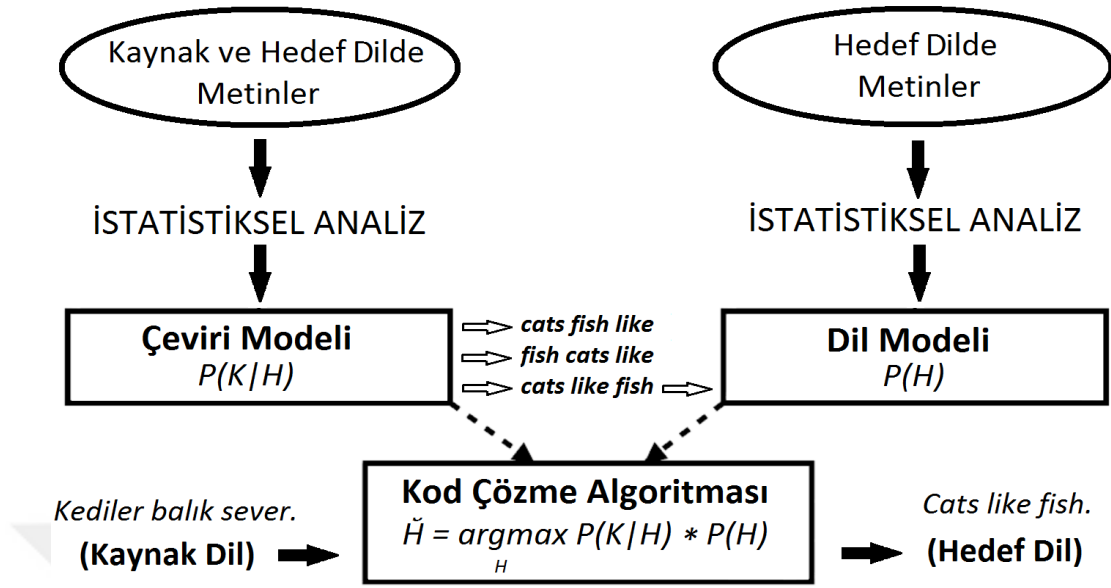
$$P(H|K) = \frac{P(K|H) P(H)}{P(K)} \quad (2.1)$$

$P(H|K)$ , kendisine kaynak dilde K sunulduğunda hedef dildeki H'yi üretme olasılığını temsil eder.  $P(H)$ , K'yi hesaba katmadan H'nin doğru dizilimini ölçen **dil modeline**;  $P(K|H)$  ise H dizisine göre K dizisinin olasılığını ölçen **çeviri modeline** işaret etmektedir. Bu iki formülü daha açık hale getirmeden önce H'den bağımsız olan paydadaki  $P(K)$ 'yi çıkararak yukarıdaki denklemi (2.1) daha basit hale

getirelim. Bu şekliyle hata ihtimalini minimuma indirmek için sistemin,  $P(K|H)$  ve  $P(H)$  olasılıklarının çarpımını maksimize etmesi yeterli olacaktır (2.2).

$$\hat{H} = \underset{H}{\operatorname{argmax}} P(K|H) * P(H) \quad (2.2)$$

Denklemden (2.2) gösterilen  $P(K|H)$  kaynak dildeki K ve hedef dildeki H cümlelerinin doğruluk olasılığı oranını temsil eder. Bu oran sıfır ile bir arasında puanlandırmayla gösterilir. Bir örnek üzerinden açıklamak gerekirse, yanlış çeviri olan “*Bugün çok iyi çalıştı.*” | “*It will be sunny tomorrow.*” K ve H cümlelerinin  $P(K|H)$  oranı sıfıra yakın puanlanırken; doğru çeviri olan “*Yarın hava güneşli olacak.*” | “*It will be sunny tomorrow.*” cümlelerinin olasılık oranı ‘1’e yakındır. Sistem her cümleyi muhtemel çeviri olarak dikkate aldıktan sonra gerçeğe en yakın çevirileri ‘1’e yakın bir değerlendirme ile öne çıkarmaktadır.  $P(H)$  ise kaynak dili hesaba katmayarak yalnız hedef dildeki kelime dizilişlerini dikkate alarak geçerli ve iyi biçimlendirilmiş bir HD cümlesinin gerçekleşme olasılığına bakar. Bunu yaparken aynı kelimelere sahip “*Cats fish like.*”, “*Fish cats like.*” ve “*Cats like fish.*” örneklerinden herhangi birini seçmek yerine, kullanım oranlarına göre en sık yan yana gelen “*Cats like fish.*” kelimelerini öne çıkarır. Bu iki  $P(H)$  ve  $P(K|H)$  formül verilerinden yararlanarak çeviri işleminin en gerçeğe yakın şekilde olup olmadığına karar veren bileşen ise **kod çözücüdür**. Koehn ve Knight’ın (2003) örneğinden esinlenerek Şekil 2.8 ile görsel hale getirilen çeviri modeli (ÇM), dil modeli (DM) ve kod çözücü (KÇ) bileşenleri, ilgili denklemleri de eklenerek sunulmuştur.



Şekil 2.8: İstatistiksel Makine Çevirisi bileşenleri

Yukarıdaki gösterimde (Şekil 2.8) ‘kaynak ve hedef dil metinler’ iki dilli KD ve HD derlemine; ‘hedef dilde metinler’ ise tek dilli HD derlemine temsil etmektedir. Çeviri modeli (ÇM) çalışma prensibini iki dilli derlem analizleri üzerinden oluştururken dil modeli (DM) tek dilli derlem üzerinden çözümlenmelerini sunar. Kod çözme algoritması ise ÇM ve DM verilerini maksimize eden bir üretim aşamasını ifade eder.

İMÇ yaklaşımının anlaşılabilmesi için gösterilen (Şekil 2.8) dört aşamayı bilmek gerekir. Birinci üzerinde durulması gereken konu derlemelerdir. Sistemde iki tür derlem vardır: Bunlar, çok büyük sayıda KD ve HD paralel metinleri bulunduran **iki dilli derlem** ve sadece HD metinleri içeren **tek dilli derlem**dir. İMÇ’lerde derlemelerin, çevirilerde referans noktaları olmaları nedeniyle anlam, şekil ve hatta amaç bakımından uygun cümleler bulundurmalarına dikkat edilmelidir. Çünkü derlemeler çeviri işleminin nasıl olacağına kaynak teşkil eden **eğitim verileridir**.

İkinci bilinmesi gereken bir **çeviri modeli** oluşturabilmek için derlemelerden faydalanan öğrenme sistemidir. Bu öğrenme sistemi sayesinde oluşturulan çeviri modeli, farklı dillerdeki kelimeler ve ifadelerin arasındaki bağlantıları tespit eder. Sistem, kelimeleri hizalayan ve çeviri olasılıklarını çıkaran yinelemeli **beklenti**

**maksimizasyonu** (*expectation maximization*) algoritmasıyla çeviri modelini öğrenir. Öğrenilen model, KD girdilerinin kelimelere ayrılması işleminden sonra her birine yeterli karşılığı (*adequacy*) olabilecek HD çeviri adaylarını sıralar. Bu kelime adaylarına, KD'nin en uygun karşılığı olan HD cümlelerine çevrilmesini sağlayacak olasılık puanlamaları atanır. İki dilli derlemdeki aramalar sırasında HD ve KD çeviri eşlerinde karşılıklı rastlanan kelimelerin, kullanım sıklığı oranlarına göre yapılmış olan bu olasılık puanlamaları çevirinin **güvenilirliği** (*faithfulness*) için gereklidir.

Üçüncü aşama HD'nin akıcılığını yansıtan **dil modelidir**. Dil modeli de çeviri modeli gibi eğitim verilerinin analizi yoluyla elde edilir ve çıktıların hedef dilde üretilme olasılığı bu model sayesinde gerçekleşir. Model, cümlelerin **akıcılığının** (*fluency*) temini için çeviri modelinden aldığı kullanım sıklık oranları verilmiş HD kelimelerini olabilecek en iyi sıralamayı gerçekleştirecek şekilde düzenler. Kelimelerin yerlerini doğru belirleyebilmek için dil modeli, tek dilli derlem içindeki HD cümlelerinin sıralanışını dikkate alan bir n-gram yönteminden yararlanır. Bu yolla dil modeli, bir dizi kelime içeren H cümlesi için P(H) formülasyonunu (2.3) kullanarak bir sözdizimsel doğruluk olasılığı ortaya çıkarabilir.

$$P(H) = P(t_1, t_2, \dots, t_n) = P(t_1)P(t_2|t_1)P(t_3|t_1, t_2) \dots P(t_n|t_1 \dots t_{n-1}) \quad (2.3)$$

Ancak bu olasılık tahmini cümledeki kelime sayısı arttıkça geçerliliğini yitirecektir. Bu yüzden sınırlı sayıda kelime üzerinden hesaplamalar tercih edilir. Bunun için genelde bir kelimenin, önündeki iki kelime ile birlikte kullanılması olasılığını tahmin eden trigram modeli (2.4) tercih edilir.

$$P(t_i|t_1 \dots t_{i-1}) \approx P(t_i|t_{i-2}t_{i-1}) \quad (2.4)$$

Dördüncü aşamada bahsedilmesi gereken bu iki çeviri modeli P(K|H) ve dil modeli P(H) formül verilerinden yararlanarak çevirinin en gerçeğe yakın şekilde üretilmesinde karar verici olan **kod çözücüdür**. Kod çözümü algoritmasındaki *argmax* operatörü (Şekil 2.8) ÇM ve DM olasılıklarını H değişkeni üzerinde en üst düzeye çıkaran bir denklemlerle çeviri işlemini gerçekleştirir. İMÇ'lerde çeviri

işlemlerinin düğüm noktası, kod çözücü algoritmasının en doğru çeviriye karar vermesi aşamasıdır. Doğru çeviri ile muğlak olmayan çeviriler kastedilmektedir.

Bir cümlede muğlaklık (*perplexity*) oranı ne kadar düşük olursa kelimeler o kadar gerçeğe yakın dizilmişlerdir. Bunun için n-gram ölçümlerinin ihtiyaç duyduğu derlem verilerinin büyüklüğü ve kalitesi çok önemlidir. Goodman, veri büyüklüğü ve kalitesinin önemini 50 bin kelimelik bir veri üzerinden unigram (2.5), bigram (2.6) ve trigram (2.7) modellerini uygulayarak; her bir n-gram modeli için elde ettiği muğlaklık değerlerine araştırmasında (Goodman, 2001) yer vermiştir. Bahsi geçen n-gram modellerine ait formüller (2.5), (2.6), (2.7) aşağıda sunulmuştur.

$$P(k_1 \dots k_n) = \prod_{i=1}^n P(k_i) \quad (2.5)$$

$$P(k_1 \dots k_n) = \prod_{i=1}^n P(k_i | k_{i-1}) \quad (2.6)$$

$$P(k_1 \dots k_n) = \prod_{i=1}^n P(k_i | k_{i-2}, k_{i-1}) \quad (2.7)$$

Yukarıdaki 2.7 numaralı denklemde gösterilen trigram modeline “*cats like fish*” cümlesi uygulandığında elde edilen 3-gram parametreleri Şekil 2.9 ile görselleştirilmiştir.

$$\begin{aligned}
P(\langle \text{sos} \rangle \langle \text{sos} \rangle \text{ cats like fish } \langle \text{eos} \rangle) &= P(\text{cats} | \langle \text{sos} \rangle, \langle \text{sos} \rangle) \\
&\quad \times P(\text{like} | \langle \text{sos} \rangle, \text{cats}) \\
&\quad \times P(\text{fish} | \text{cats}, \text{like}) \\
&\quad \times P(\langle \text{eos} \rangle | \text{like}, \text{fish}) \\
k_{-1} = k_0 &= \langle \text{sos} \rangle \\
k_{n+1} &= \langle \text{eos} \rangle
\end{aligned}$$

**Şekil 2.9: Örnek cümle üzerinden trigram parametreleri**

Şekil 2.9 ile dört adet olasılık değerlerine sahip parametre sunulmaktadır. Trigram modelini her parametrede temsil edebilmek için girdi cümlesine iki adet boşluk ( $\langle \text{sos} \rangle$ ) ve bir adet cümle sonu ( $\langle \text{eos} \rangle$ ) eklenmiştir. Model bu olasılık tahminlerini büyük oranda HD cümleler ihtiva eden derlemlerden yararlanarak yapmaktadır. Örneğin, yukarıda (Şekil 2.9) gösterilen  $P(\text{fish} | \text{cats}, \text{like})$  3-gram parametresinde ‘cats like’ kelimelerinden sonra ‘fish’ kelimesinin gelme olasılığı milyonlarca kelimelelik metin analizi sonunda elde edilir. Azami Benzerlik Tahmini (*Maximum Likelihood Estimate*) olarak da ifade edilebilecek bu ölçüm aşağıda (2.8) verilmiş ve kelimeler üzerinden örneklendirilmiştir (2.9).

$$P(k_i | k_{i-2}, k_{i-1}) = \frac{\text{Topla}(k_{i-2}, k_{i-1}, k_i)}{\text{Topla}(k_{i-2}, k_{i-1})} \quad (2.8)$$

$$P(\text{fish} | \text{cats}, \text{like}) = \frac{\text{Topla}(\text{cats}, \text{like}, \text{fish})}{\text{Topla}(\text{cats}, \text{like})} \quad (2.9)$$

Azami benzerlik tahmini, denklemdeki (2.9) ‘cats like fish’ üçlüsünün görülme sayısının, ‘cats like’ ikilisinin görülme sayılarına bölünmesiyle gerçekleşen bir puanlama işlemidir. Tüm parametreler aynı işlemde geçtikten sonra cümlenin olasılık oranı belirlenmiş olur (Şekil 2.10). En uygun cümlenin hangisi olduğuna dair tahmin için gereken veriyi dil modeli, bu yolla elde eder.

Dil modeli sistemde kilit bir vazife icra etmektedir. Çünkü çeviri modelinden gelen en yüksek olasılığa sahip kelime tercümeleri cümle boyutuna indirildiğinde gerçeğe uygun HD çevirileri olmayabilir. “Yüz lira verdi.” cümlesindeki ‘yüz’ kelimesi için çeviri modeli, ‘100’ anlamına gelen ‘*hundred*’ çevirisini düşük olasılıklı gösterirken; ‘surat’ manasına gelen ‘*face*’ çevirisini ise daha sık rastlanan olduğu için en yüksek çeviri ihtimali olarak sunacaktır. Ancak dil modeli tek dilli derlemden öğrendiği bilgi ile HD’deki kelime sıralanışlarını dikkate alan bir değerlendirme yapacağı için, ‘surat’ anlamına gelen ‘*face*’ kelimesini seçmek yerine; daha düşük ihtimalle verilen ‘*hundred*’ kelimesini kullanmayı tercih edecektir. Böylece en iyi sonucu bulmakla sorumlu olan kod çözücü, çeviri modelinin sunduğu  $P(K|H)$  ve dil modelinin sunduğu  $P(H)$  olasılık analizi puanlarını en yüksek düzeye çıkararak bir formül (2.2) üzerinden çeviri görevini yerine getirmiş olacaktır. Bahsi geçen olasılık analizi puanları üzerinden en iyi çeviri tercihinin yapılması süreci, Koehn ve Knight’ın (2003) İspanyolcadan İngilizceye çeviri örneğiyle (Şekil 2.10) daha anlaşılır hale getirilebilir.

Que hambre tengo yo		
↓		
What hunger have	$p(s/e)$	= 0.000014
Hungry I am so	$p(s/e)$	= 0.000001
I am so hungry	$p(s/e)$	= 0.0000015
Have I that hunger	$p(s/e)$	= 0.000020 ←

**Şekil 2.10: İspanyolca-İngilizce ÇM olasılık puanları**

İspanyolca “Que hambre tengo yo” KD girdisinin dört aday HD çevirileri ile her çevirinin yanında çeviri modeli değerleri olan  $P(K|H)$  oranları verilmiştir (Şekil 2.10). Burada sadece çeviri modeli verileri üzerinden bir analiz yoluna gidildiğinde görselde de işaret edilen ‘*Have I that hunger*’ cümlesinin en yüksek değere sahip olduğu için doğru çeviri olduğu düşünülebilir; ancak adaylardan hangisinin daha gerçeğe yakın çeviri olduğunun anlaşılabilmesi için dil modeli puanlamalarının da hesaba katılması gerekir (Şekil 2.11).

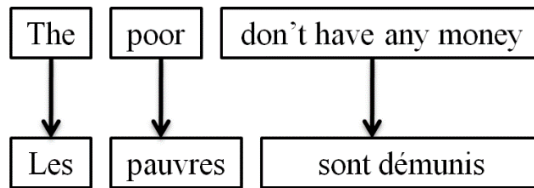
Que hambre tengo yo	
↓	
What hunger have	$p(s/e) p(e) = 0.000014 * 0.000001$
Hungry I am so	$p(s/e) p(e) = 0.000001 * 0.0000014$
I am so hungry	$p(s/e) p(e) = 0.0000015 * 0.0001$ ←
Have I that hunger	$p(s/e) p(e) = 0.000020 * 0.00000098$

**Şekil 2.11: İspanyolca-İngilizce ÇM ve DM olasılık puanları**

Görüldüğü gibi Koehn ve Knight’ın (2003) İspanyolca-İngilizce çeviri örneğinde  $p(s/e)$  çeviri modeli değerleri ile  $p(e)$  dil modeli ölçümlerinin çarpılması sonucunda en yüksek olasılık değerini “Que hambre tengo yo.” KD cümlesinin “Have I that hunger” yanlış HD çevirisi (Şekil 2.10) yerine; daha gerçeğe yakın olan “I am so hungry” cümlesi (Şekil 2.11) almıştır.

### 2.5.2.1. Kelime tabanlı İMÇ’den ifade tabanlı İMÇ’ye

İlk kez 1993 yılında tasarlanan kelime tabanlı İMÇ’ler (IBM modelleri vb.) KD bir kelime için 0, 1 veya n sayıda HD kelime olacak şekilde karşılık tanımlayabilir. Örneğin, ‘ev’ kelimesini ‘the house’, ‘bozuk’ kelimesini ‘out of order’ ile hizalar (*alignment*). Ancak tersi durum olan, birden fazla KD kelimenin HD kelime veya kelimelerle hizalanması (Şekil 2.12) örneği (Brown vd., 1990) durumlarında kelime tabanlı sistemler yetersiz kalır.

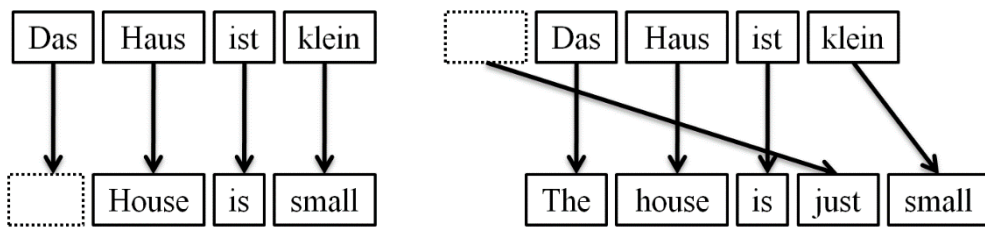


**Şekil 2.12: İMÇ’lerde birden fazla kelimeli yapıların hizalanması**

Burada “don’t have any money” dizilimi Fransızca "sont démunis" grubuna karşılık gelmektedir. Bu dört İngilizce kelimenin, ki ‘don’t’ (*do not*) kelimesini tek kelime kabul ederek, iki Fransızca kelimeye karşılık gelmesi bir m-n karşılaştırma durumuna (m=4, n=2) örnektir (Poibeau, 2017). HD’deki birleşik bir kelime grubunun KD’deki bir birleşik kelime grubu ile hizalanması beklenemez. Çünkü kelimeyi çeviri metodunun temeli gören bir yaklaşım için cümlenin daha büyük bir yapısı olan ifade (*phrase*) yapılarını birleştirebilmesi mümkün değildir.

Bazen hayattan alınan bir bilgi de çeviriyi geçersiz hale getirir. “*I took a train to Berlin.*” cümlesi, “*I took a train to Paris.*” cümlesi olarak çevrilebilir. Bu ilginç durumun sebebi eğitim verisinde büyük oranla ‘*train to Paris*’ kelmelerinin yan yana görülmüş olmasıdır (Okpor, 2014). Eklemek gerekir ki aynı çevirinin Türkçeye yapılması durumunda sadece veri sorunu değil; bazı kelimelerin çıkarılması durumuyla da karşılaşılır. Örneğin; ‘*a train to Berlin*’ = ‘~~bir~~ Berlin treni’ gibi.

Kelimelerin eklenmesi veya çıkarılması gibi durumların kelime tabanlı sistemlerde karşılaşılan en büyük zorluklardan biri olduğu görülmüştür. Koehn’un (2007) bir ders örneğinden alınan kelime ekleme veya kelime çıkartma durumlarına işaret eden görseli Şekil 2.13 ile sunulmuştur.



**Şekil 2.13: İMÇ’lerde kelime ekleme-çıkarma durumlarında hizalama sorunu**

“Das Haus ist klein” cümlesinde birinci örnekte HD için boş (*null*) kelime üretilmişken; ikinci örnekte KD cümlesinde kullanılmamış bir kelime HD’de ‘*just*’ olarak üretilmiştir.

Kelime tabanlı yaklaşımların bir diğer eksikliği çeviri sırasında bağlam bilgisinin olmamasıdır. Genelde kelimeler gruplar halinde çevrilme eğilimindedir ve kelime kelime çeviri her zaman bütün bir cümlenin gerçek anlamını veremez. Herhangi bir kelime için hedef dildeki çeviri ve konum, sınırlandırma etkisi (localization effect) olarak da adlandırılan, etrafındaki kelimelere bağlı olarak farklılık gösterebilir. Örneğin 'quit' fiili 'quit smoking' bağlamında 'bırakmak' olarak çevrilirken; aynı kelime 'quit the program' için 'çıkmaq' olarak çevrilir. Ancak Kelime Tabanlı model bu gibi durumlarda yalnız dil modelini kullanıp çeviri modelini süreç katamadığı için yetersiz kalır (Durgar El-Kahlout, 2009). Bu ve benzeri sebeplerle Kelime tabanlı sistemler neredeyse tamamen terk edilmiş ve yerlerini ifade tabanlı çevirilere bırakmışlardır (Skadins, 2011).

#### 2.5.2.2. İfade tabanlı İMÇ

İfade tabanlı çeviri modellerinde gürültülü kanal (*Noisy Channel*) yaklaşımı kullanılabilir; ancak ifade tabanlı çeviri modelleri, ifade çeviri olasılıklarını biçimlendirmek için **log doğrusal modelleri** (*Log-linear Model*) kullanma eğilimindedir (Koehn vd, 2003). Bu yaklaşım P (t|e)'yi özellik fonksiyonlarının ağırlıklı bir kombinasyonu olarak modeller. Bu yaklaşım dil ve çeviri modellerine ekstra özelliklerle güç kazandırılmak istendiğinde kullanılan, gürültülü kanal modelinin geliştirilmiş versiyonudur (Durgar El-Kahlout, 2009). Makine Çevirisinin temel denklemi (2.2) yalnızca iki özneliği barındırır: P(K|H) ve P(H). Ancak çeviri kalitesi için başka özelliklere de ihtiyaç duyulabilir. Örneğin, çevrilen cümlelerin uzunluğunun süreç içinde değerlendiriliyor olması faydalı bir öznelik sayılabilir. Koehn vd (2003) tarafından tarif edilen ifade tabanlı log-linear model en makul çeviriyi elde edebilmek için bazı farklı özellikler sunmaktadır (Skadins, 2011).

SMT sistemlerinin çoğunda,  $f$  kaynak cümlesinin verildiği bir çeviri hipotezi  $e$  olasılığı, log-doğrusal bir model (2.10) olarak formüle edilir:

$$P(e|f) = \frac{1}{Z(f)} \exp\left(\sum_{k=1}^K \lambda_k h_k(e, f)\right) \quad (2.10)$$

Denklemden (2.10) modeldeki öznitelik sayısı  $K$ , öznitelik ağırlıkları  $\lambda_k$  ve öznitelik fonksiyonları  $h_k$  ile gösterilmiştir.  $Z$  aslında bir normalizasyon faktörü olduğu için  $\sum_e P(e|f)=1$  olacaktır. En iyi  $e$  hipotezi için karar kuralı,  $P$  üzerinde doğrudan bir maksimizasyondur (2.11), (2.12).

$$e^* = \arg \max_e P(e|f) \quad (2.11)$$

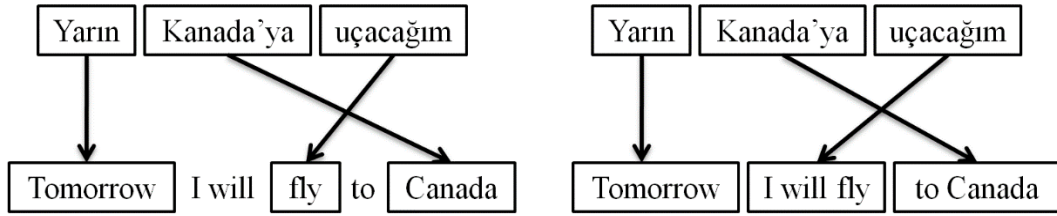
$$= \arg \max_e \sum_{k=1}^K \lambda_k h_k(e, f) \quad (2.12)$$

Bu sistemlerde genellikle  $K$  ile birlikte 10-15 civarı  $h_k$  öznitelik kullanılmaktadır. En sık kullanılan öznitelikleri şöyle sıralanabilir: Her iki çeviri yönünde kelime çevirme olasılıkları (*word translation probabilities*); her iki çeviri yönünde ifade çevirisi olasılıkları (*phrase translation probabilities*) (sistemden sisteme değişir); hipotezin dil modeli olasılığı ( $h_{LM}(e, f) = P(e)$ ); hipotezdeki kelime / ifade / kural sayısı; yeniden sıralama (*distortion*) model maliyeti vb. Bu sıralananların çoğu, kod çözme öncesi gerçekleşen öğrenim aşamasında eğitim verileri olan derlemler üzerinden hesaplanması gereken olasılıklardır (Mermer, 2019).

Birleşik olasılık yöntemi (*joint probability method*) kullanan bir yaklaşım olan ifade tabanlı çeviri modeli ilk kez Marcu ve Wong (2002) tarafından sunulmuştur. Yaklaşım daha sonra Koehn vd (2003) tarafından revize edilmiştir. Kelimeyi çeviri metodunun temeli olmaktan çıkararak yaklaşım cümlelerin daha büyük bir yapısı olan ifade (*phrase*) yapılarını kullanan bir metod izler. Yani HD’de birleşik bir kelime grubu, KD’deki bir birleşik kelime grubu ile hizalanabilir. Böylece cümle içinde birkaç kelimenin kendi aralarında oluşturdukları manalar da çeviri sürecinde dikkate alınmış olur. Yani KD veya HD aralarında oluşabilecek local kelime grupları da

çevrilebilir hale gelir (Ahmed ve Hanneman, t.y.) Kelime olarak ‘ifade’ tabiri cümle içinde yan yana olduklarında farklı anlam kazanan kelime dizileri için kullanılmaktadır. Burada, isim grubu (*noun phrase*) veya fiil grubu (*verb phrase*) gibi dilbilimsel bir ifade kastedilmemektedir.

İMÇ’lerde çevirinin nasıl gerçekleşeceği, sürecin en başındaki eğitim verilerinin nasıl öğrenildiğiyle bağlantılıdır. Yani sistem yalnız kelimeleri baz almayı en az iki kelimedenden oluşan ifade yapılarını da öğrenme sürecine katarsa, kelime tabanlı sistemlerde görülen eksiklikler önemli ölçüde giderilmiş olur. Veri ihtiyacı ve işlem hızı yönünden maliyetli görünse de bu, sonuçları itibariyle ifade tabanlı yöntemlerin kelime tabanlı yöntemlere olan tercih üstünlüğünü değiştirmez. Aradaki farka işaret edebilmek için Durgar El-Kahlout’un (2009) her iki yöntemde de sunmuş olduğu İngilizce-Türkçe çeviri örneklerini (Şekil 2.14) incelemek yeterli olacaktır.



**Şekil 2.14: Kelime tabanlı ve ifade tabanlı çeviri karşılaştırması**

Şekil 2.14 ile sunulan birinci örnekte yalnız kelimeyi temel alan bir çözümleme yolu takip edilerek ‘uçacağım’ kelimesi sadece ‘fly’ kelimesiyle; ‘Kanada’ya’ kelimesi de aynı şekilde yalnız ‘Canada’ kelimesiyle eşleşmiştir. Ancak verilen ikinci örnekteki (Şekil 2.14) hizalama yöntemi, ifade boyutunda gerçekleştirildiği; yani kelimeleri bir bütün olarak ele aldığı için en doğru çeviriyi gerçekleştirebilmektedir.

Ancak sistemin öğrenme yaklaşımının, anlamsal açıdan her zaman iyi sonuçlar vereceğini söylemek fazla iyimserlik olur. Çünkü sistemin çalışma prensibi istatistiği kullanarak KD bir ifadenin HD bir ifadeyle yer değiştirilmesine dayanır. Eğer ifadelerden biri eğitim verisinde daha çok kullanılmışsa, yanlış çeviri de olsa dil

modeli tarafından doğru çeviriye tercih edilecektir (Alsohybe, 2017). Burada çarpıcı bir benzetme olarak İngiliz devlet adamı Benjamin Disraeli'nin sözü hatra gelmektedir: “Üç çeşit yalan vardır; basit yalan, kuyruklu yalan ve istatistik.”

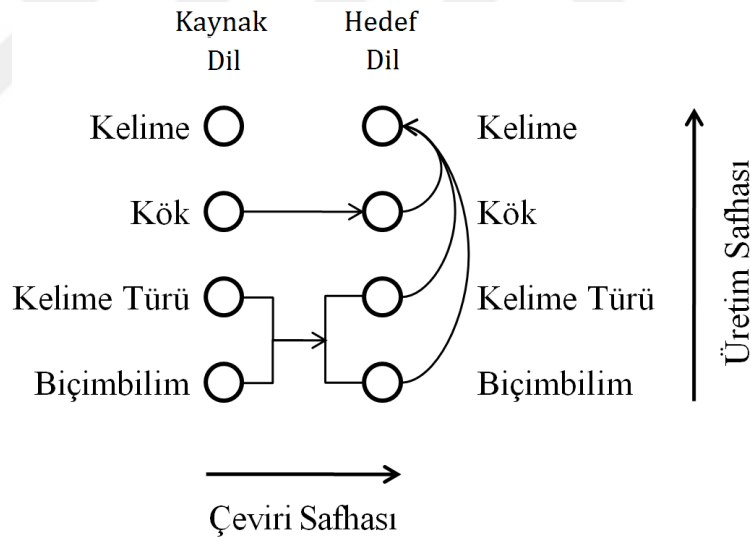
### 2.5.2.3. Faktörlü İMÇ'ler

İfade Temelli İMÇ'ler morfolojik, sözdizimsel veya anlamsal olsun; dilbilimsel motivasyonlarla tasarlanmış değildir. Ancak doğan ihtiyaçlar gereği dilbilgisi kurallarının da sisteme eklendiği bir çok model geliştirilmiştir. Dilbilimsel ayrıştırmaları kaynak cümle seviyesinde eklemleyen (Huang vd., 2006); hedef cümle seviyesinde uygulayan (Galley et al., 2006) veya hem kaynak hem de hedef dilde bu dilbilimsel ayrışmalardan yararlanmayı tercih eden (Zhang ve Gildea, 2008) sistemler mevcuttur. Ancak, dilbilgisinin çeviri yaklaşımlarına daha sıkı entegre edilmesi temelde iki nedenle arzu edilmektedir. Birincisi kelimeleri yüzeysel biçimlerinden çok köken bilgisi (*lemmas*) gibi daha genel temsiller üzerinden ele alan çeviri yaklaşımlarının daha zengin istatistikler ortaya koyacağı ve bu yolla sınırlı eğitim verilerinden kaynaklanan veri sıkıntısının üstesinden gelinebileceği düşüncesidir. İkinci neden ise çeviride birçok durumun biçimbilimsel, sözdizimsel veya anlambilimsel düzeyde daha iyi açıklanabiliyor olmasıdır. Tüm bunlardan faydalanmak çevirinin direkt modellenemesine yardımcı olur. Örneğin; cümle seviyesinde yeniden sıralama, (*reordering*) kelimelerin bulunduğu yere (sözdizime) bağlı iken kelimelerin yeri ise aldıkları biçimlere (biçimbilime) bağlıdır (Koehn, 2007).

Bu yeni yaklaşım ek bilgi notlarına kelime seviyesinde izin vermektedir. Bu yapı içerisinde kelime yalnız bir işaret değil, farklı düzeylerde bilgi notu içeren unsur yüklenicidir. (Koehn, 2007). Bu özellik sayesinde bir kelime önceki istatistiksel yöntemlerdeki gibi izole halde değildir. Yani sistemde, aynı fiil köküne sahip olup çekime uğramış veya aynı isim köküne sahip olup türetilmiş iki kelimenin biçimbilimsel yakınlıklarına rağmen farklı formlarda olmaları nedeniyle farklı kelimeler gibi algılanmaları durumunun önüne geçilmiştir. Örneğin; ‘kedi’ kelimesi eğitim verisinde bulunmuyor olsa da, artık ‘kediler’ kelimesinden bağımsız bir

kelime olarak görülmez. Sıradan bir İMÇ modelinde kelimeler tamamen farklı iki kelime olarak görülür.

Dilbilimsel bilginin desteğini yanına alarak araştırmacıların ilgisi çeken faktörlü çeviri modelleri (*factored translation models*) (Koehn vd., 2006; Avramidis vd., 2008) aslında morfolojik açıdan zengin diller için kullanılmak üzere tasarlanmıştır. Arapça, Japonca veya Türkçenin tek bir kökle birçok farklı kelime formu türetebilme özelliğine sahip olması, bu diller için özellikle sınırlı eğitim verilerinde sağlıklı istatistiklere ulaşabilmeyi zorlaştırmaktadır. Bu gibi durumlar için faktörlü çeviri, kökleri ve biçimleri ayrı ayrı çeviren ve ardından hedef yüzey formunu oluşturan daha genel bir yaklaşım sunar (Yeniterzi, 2009). Şekil 2.15 ile bu model, görsel hale getirilmiştir.



**Şekil 2.15: İMÇ'lerde Faktörlü çeviri modeli**

Verilen örnekteki (Şekil 2.15) oklar eşleştirme safhalarını temsil etmektedir. Bu modelde İki tür eşleştirme safhası vardır: (i) İlki, girdi faktörlerini ifade düzeyinde çıktı faktörleriyle eşleyen çeviri safhasıdır. Çeviri safhası altta sağa yönlü çizilmiş ok ile gösterilmektedir. Bu modelde iki çeşit çeviri adımı vardır: (a) Giriş köklerinin çıkış köklerine çevrilmesi ve (b) kelime türü ile biçimbilim bilgisi girdilerinin, kelime türü ve biçim bilgisi çıktılarına dönüştürülmesidir. (ii) Diğer eşleştirme ise

üretim safhasıdır. Bu adımda çıktı faktörlerinin diğer çıktı faktörleriyle kelime düzeyinde eşleştirilmesi yapılır. Kök, kelime türü ve biçimbilimle yüzey formunun oluşturulması eğimli dikey oklarla temsil edilmektedir (Yeniterzi, 2009). Örnekte (Şekil 2.15) KD'deki kökün, HD'deki kökle; kelime türü ve biçimbilimsel bilginin, HD'deki kelime türü ve biçimbilimsel bilgi ile eşleştiği görülmektedir. Sonrasında HD'deki bu faktörler, yüzeysel biçimi üretimi (*generation*) aşamasında kullanılır. Böylece kök, kelime türü ve biçimbilimsel bilgidен yüzeysel bir biçim oluşturulmuş olur (Yıldırım, 2014).

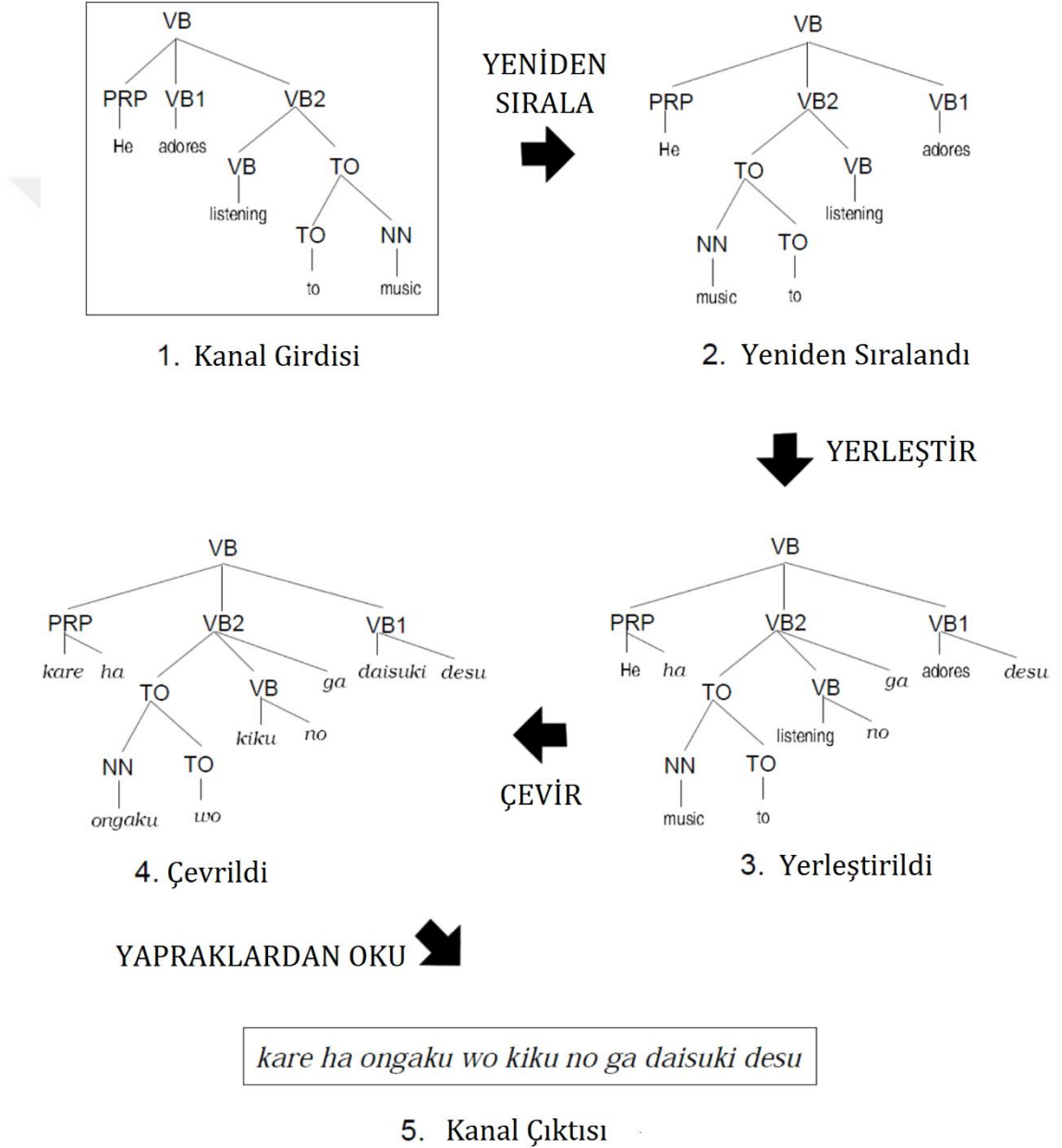
Çeviri sistemlerindeki işlem ve veri problemleri, teoride mümkün olan ancak gerçekleştirilemeyecek kadar maliyetli ve karmaşık faktörlü modellerin geliştirilmesine imkân tanımamıştır (Koehn ve Hoang, 2007). Burada, tasarım kaygıları noktasında ortak oldukları söylenebilecek başka çeviri modellerinden de bahsetmek yerinde olacaktır. Bunlar sözdizim tabanlı ve hiyerarşik ifade tabanlı çeviri modelleridir. Bu modeller de faktörlü çeviri modeli gibi morfolojik zenginliğe sahip veya uzak yapıya sahip olan doğal dilleri dikkate alarak geliştirilmişlerdir.

#### **2.5.2.4. Sözdizim tabanlı İMÇ'ler:**

Yamada ve Knight (2001) sözdizimi tabanlı istatistiksel bir çeviri modeli sunmuştur. Modelde her düğümde stokastik işlemler uygulanarak bir KD ayrıştırma ağacı HD dizgisine dönüştürülür. İşlem sırasında kelime sırası ve hal durumları (bulunma, yönelme, vb.) gibi dil özellikleri hesaba katılır. Model parametreleri, EM (*Expectation Maksimation*) algoritması kullanılarak polinomial zamanda tahmin edilir. Modelin kelime hizalaması IBM Model 5 tarafından üretilenlerden daha iyidir. Yamada ve Knight IBM tarzı ÇM'yi, dilin yapısal veya sözdizimsel yönlerini modellememesi noktasında eleştirmiş; İngilizce ve Japonca gibi çok farklı kelime sıralamasına sahip bir dil çiftinin bu ÇM'lerce iyi modellenmeyeceğini ileri sürmüştür.

Model, dilin yapısal yönlerini dâhil etmek için girdi olarak ayrıştırma ağaçlarını kabul eder. Yani giriş cümlesi bir sözdizim ayrıştırıcı tarafından önceden işlenir. Kanal, ayrıştırma ağacının her bir düğümü üzerinde işlemler gerçekleştirir. İşlemler, alt düğümleri **yeniden sıralamak**, her düğüme kelimeler **yerleştirmek** ve yaprak

düğüm kelimelerini **çevirmek** şeklinde gerçekleşir. Şekil 2.16 ile bu üç aşama görsel hale getirilmiştir. Modelin çıktısının ayrıştırma ağacı değil, bir dizge olduğu fark edilmektedir. Ayrıştırma yalnız kanal girişi tarafında gereklidir (Yamada ve Knight, 2001).



Şekil 2.16: Sözdizim tabanlı İMÇ'lerde işlem aşamaları

Şekil 2.16 ile gösterilen modelde kanal girdisi, önce KD’de ayrıştırma ağacı haline getirilmiştir. Ağaç düğümleri HD kelimeler ile eşleştirilmesi öncesi HD sözdizim yapısı dikkate alınarak sıralanması işlemine tabi tutulur. Modeldeki **yeniden sıralama** işlemi, ÖYN (özne+yüklem+nesne) dilleri (İngilizce veya Çince) ve ÖNY (özne+nesne+yüklem) dilleri (Japonca veya Türkçe) gibi farklı kelime sıralamalarına sahip diller arasındaki çeviriyi modelleyebilmek için tasarlanmıştır. Kelime **yerleştirme** işlemi sözdizimsel durumları belirlemede dil farklılıklarını yakalamaya yöneliktir. Örneğin, İngilizce ve Fransızcada bulunma, yönelme, ayrılma gibi hal durumları için kelimenin cümle içindeki konumu etkiliyken; Japonca ve Korece gibi dillerde aynı fonksiyonlar için kelimeye hal ekleri eklenir. Ekstra kelimeler düğümlere stokastik olarak yerleştirilir. Kelimenin düğümün sağına ya da soluna yerleştirilmesi durumu girdi ayrıştırma ağacına göre belirlenir. Yerleştirilen HD kelimeler KD kelimeler ile değiştirilir ve son olarak tüm yapraklardaki kelimelerin okunması sağlanarak kanal çıktısı elde edilmiş olur (Yamada ve Knight, 2001).

Sözdizim tabanlı istatistiksel modeller uzak yapılaraya sahip dillerin çevirileri için tasarlanmıştır. ÖYN (Özne+Yüklem+Nesne) yapısına sahip İngilizce ile ÖNY (Özne+Nesne+Yüklem) yapısına sahip Türkçe çevirilerinde bu tür modeller tercih edilebilir. Sözdizim tabanlı İMÇ’ler gibi uzak yapıya sahip dillerin çevrilmesinde kullanılan bir başka model ise hiyerarşik ifade tabanlı İMÇ’lerdir.

#### **2.5.2.5. Hiyerarşik ifade tabanlı İMÇ:**

Chiang (2007) mevcut son teknoloji MÇ sistemleri ile çözülemeyen tipik sorunlara değinmiş ve hiyerarşik ifadeler (alt ifadelerden oluşan ifadeler) kullanan istatistiksel bir makine çevirisi modeli ortaya koymuştur. Hiero adlı bu model, Aho ve Ullman’ın (1969) ‘ağırlıklı senkronize bağlamdan bağımsız gramer’ini (*Weighted synchronous Context-free grammar*) temel alır. Sistem, biçim itibarıyla ‘sözdizim yönlendirmeli dönüştürme grameri’ (*syntax-directed transduction grammar*[Lewis ve Stearns, 1968]) olarak da adlandırılan ‘eş zamanlı bağlamdan bağımsız gramer’dir ancak öğrenme işlemi herhangi bir sözdizimsel bilgi notları içermeyen paralel metinlerden yapar. Bu nedenle temel fikir olarak hem sözdizim hem de ifade tabanlı çevirinin birleşimi olarak görülebilir. Chiang, bu modelin; son teknoloji hizalama

şablonu sistemi olan ifade tabanlı modellerden önemli ölçüde daha iyi performans gösterdiğini iddia etmektedir. Ona göre diller ifadelerden oluştuğu ve bu ifadeler hiyerarşik bir şekilde sıralandığı için odaklanılması gereken nokta, uzun mesafeli yeniden sıralama sorunudur.

Chiang'a (2005) göre ifadeler, kelimenin yeniden sıralanmasında işe yaradığı gibi ifadenin de yeniden sıralanmasında işe yarayabilir. Bunun yapılabilmesi için, içinde ifadeleri barındıran **hiyerarşik ifadeler**e ihtiyaç vardır. Bu hiyerarşik ifadeler için aşağıda bir örnek (2.13) sunulmuştur.

$$\langle \text{have } \boxed{1} \text{ with } \boxed{2}, \boxed{2} \text{ ile } \boxed{1}\text{-e sahip} \rangle \quad (2.13)$$

Kutu içinde gösterilen  $\boxed{1}$  ile  $\boxed{2}$  rakamları, alt ifadeler için; yani ifade içindeki ifadeler için yer tutucu görevi ifa etmektedir (Chiang, 2005). Yukarıdaki örneğin (2.13) solundaki İngilizce ifadelerden “with  $\boxed{2}$ ” edat grubu (*prepositional phrase*) sonda; sağdaki Türkçe ifadelerden “ $\boxed{2}$  ile” edat grubu ise başta. Bu edat gruplarının iki dilde de bu şekilde gözlenme sayılarının artması, sistemin bunu bir kural olarak tanımlaması için yeterlidir.

Kuralların öğrenilmesi sırasında hiç bir dilbilimsel varsayım veya bilgi notu desteği olmaksızın yalnız paralel textlerden yararlanılır. Öğrenilen kurallara göre ifade çiftleri, belli bir sıra ile birleştirilir. Bu birleşen ifadelerin meydana getirdiği hiyerarşik yapı kendisinden daha üstteki bir ifade ile birleşerek büyür. Hiyerarşik ifade çiftlerinin, senkronize bağlamdan bağımsız gramer (*synchronous CFG*) kuralı olarak biçimlendirilmiş hallerine üç adet örnek (2.14), (2.15) ve (2.16) ile verilmiştir.

$$X \rightarrow \langle \text{have } X_2 \text{ with } X_1, X_1 \text{ ile } X_2\text{-e sahip} \rangle \quad (2.14)$$

$$X \rightarrow \langle \text{the } X_2 \text{ that } X_1, X_1 \text{ olan } X_2 \rangle \quad (2.15)$$

$$X \rightarrow \langle \text{one of } X_1, X_1\text{-den biri} \rangle \quad (2.16)$$

'X' cümle içinde bir kelime veya ifadeye karşı düşmeyen ama ifadelerin hangi seviyede tutulduğunu gösteren bağlantı ucudur (*nonterminal symbol*). Çünkü bu model, ifadelere sözdizimsel kategoriler atamak yerine; onları bağlantı uçları ve indisler yoluyla belli konumlara yerleştirir. Bağlantı uçlarındaki indisler çevrilmesi gereken ikili ifadelerin birbiriyle bağlantısına ve hangi seviyede olduklarına işaret eder. Kurala göre birleşmesi gereken iki 'X' (bağlantı ucu) yeni bir X altında birleşir. Birleşen X'ler artık bu yeni indisli X ile temsil edilir. Bu birleşme işlemi tüm kurala göre eşleşme sırası gelen X'lerin ayrı ayrı x'ler halinde yan yana gelmelerine kadar devam eder (Şekil 2.17). Sistemde tüm kurallar X ile gösterilirken S ile ifade edilen yapışma kuralı (*glue rule*) buna dahil edilmemiştir. S ile gösterilen bu birleştirme uçları yardımıyla en son yan yana sıralanmış olan X'lerin birleştirilmesi sağlanır. S formunda bu görevi yerine getiren iki adet yapışma kuralı aşağıda (2.17) ve (2.18) ile gösterilmiştir (Chiang, 2005).

$$S \rightarrow \langle S_1 X_2, S_1 X_2 \rangle \quad (2.17)$$

$$S \rightarrow \langle X_1, X_1 \rangle \quad (2.18)$$

Yapışma kuralı sayesinde X ifadelerinin çevirileri ayrı ayrı gerçekleşebilmekte ve sonrasında standart ifade tabanlı (phrase-based) modeller gibi birleşmeleri sağlanmaktadır (Chiang, 2007). Şekil 2.17 ile senkronize bağlamdan bağımsız gramer (*synchronous context-free grammar: SCFG*) yoluyla cümle çevirisi oluşturulması süreci görsel hale getirilmiştir.

< S <sub>1</sub> , S <sub>1</sub> >
14 → < S <sub>2</sub> X <sub>3</sub> , S <sub>2</sub> X <sub>5</sub> >
14 → < S <sub>4</sub> X <sub>5</sub> X <sub>3</sub> , S <sub>4</sub> X <sub>3</sub> X <sub>5</sub> >
15 → < X <sub>6</sub> X <sub>5</sub> X <sub>3</sub> , X <sub>6</sub> X <sub>3</sub> X <sub>5</sub> >
9 → < Bursa X <sub>5</sub> X <sub>3</sub> , Bursa X <sub>3</sub> X <sub>5</sub> >
11 → < Bursa is X <sub>3</sub> , Bursa X <sub>3</sub> dir>
8 → < Bursa is one of X <sub>7</sub> , Bursa X <sub>7</sub> den biridir>
7 → < Bursa is one of the X <sub>9</sub> that X <sub>8</sub> , Bursa X <sub>8</sub> olan X <sub>9</sub> den biridir>
6 → < Bursa is one of the X <sub>9</sub> that have X <sub>2</sub> with X <sub>1</sub> , Bursa X <sub>1</sub> li X <sub>2</sub> e sahip olan X <sub>9</sub> den biridir>
10 → < Bursa is one of the X <sub>9</sub> that have X <sub>2</sub> with traditional minarets, Bursa geleneksel minareli X <sub>2</sub> e sahip olan X <sub>9</sub> den biridir>
12 → < Bursa is one of the X <sub>9</sub> that have big mosques with traditional minarets, Bursa geleneksel minareli büyük camilere sahip olan X <sub>9</sub> den biridir>
13 → < Bursa is one of the most beautiful cities that have big mosques with traditional minarets, Bursa geleneksel minareli büyük camilere sahip olan en güzel şehirlerden biridir>

**Şekil 2.17: Senkronize bağlamdan bağımsız gramer modeli çeviri örneği**

Cümlelerin yanlarında her adımda uygulanmış olan kuralları temsilen numaralandırılmış oklar gösterilmektedir (Şekil 2.17). Bir kural uygulanıp cümlede gerekli çıkarma ekleme yapıldıktan sonra bir sonraki kuralın uygulanmasına geçilir. Cümleye tek bir seferde çözümleme uygulanmayıp bu şekilde sıra gözetilmesinin sebebi dil ifadelerinin aralarındaki hiyerarşik yapıdır. Örneğin; 6 numaralı <have X<sub>2</sub> with X<sub>1</sub>, X<sub>1</sub> ile X<sub>2</sub> -e sahip> kurallarının uygulandığı cümlede İngilizce ‘*most beautiful cities*’ ve Türkçe ‘en güzel şehirler’ ifadeleri kaldırılarak her iki cümlede de yerlerine X<sub>9</sub> bağlantı ucu (*nonterminal symbol*) yerleştirilmiştir. Ardından uygulanan 7 numaralı kuralda da aynı işlem (‘X<sub>9</sub> that X<sub>8</sub>’ İngilizce ifadesi ile ‘X<sub>8</sub> olan X<sub>9</sub>’ Türkçe ifadeleri yerine X<sub>7</sub> bağlantı ucunun getirilmesi) gerçekleştirilerek bir sonraki işlem adımıyla devam edilmiştir. 15’inci kuralın uygulanmış olduğu safhada S formu üzerinden yapışma kuralı (*Glue rule*) aşamasına geçilmiş ve kalan parçaların da ifade tabanlı sistem mantığına göre birleştirilmeleri sağlanmıştır (Şekil 2.17).

Düz yapılardan, bu şekil bir hiyerarşik yapıya geçmiş olmak çeviri kalitesinde dikkate değer bir kalite sağlamakla birlikte biçimbilimsel fikirlere olan ihtiyacı ortadan kaldırmamıştır. Mevcut model her türlü geliştirilmeye açıktır. Modelin gelecekteki çalışmalarda dilbilim destekli bir derlem tabanlı çeviri modeli olmaya yatkın olduğu söylenebilir.

Derlem tabanlı tüm çeviri modelleri kendilerinden önceki çalışmaların kapatamadığı açıkları kapatabilme motivasyonlarıyla geliştirilmişlerdir. Hutchins'in (2005) en basit tarifine göre İMÇ, doğrudan çeviri (*direct translation*) yönteminin karmaşık istatistiklerden geçirilmiş halidir.

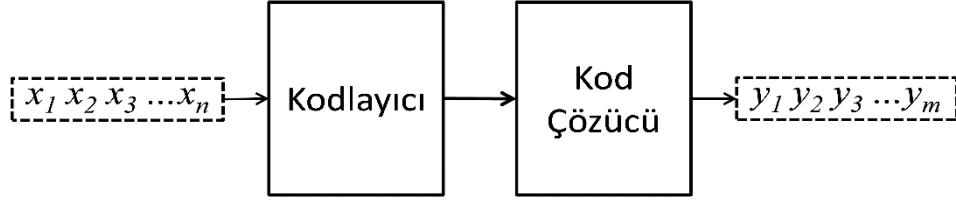
Yıllar süren gelişmeler sonunda bugün tüm büyük çalışmalar derin öğrenme adı verilen yeni bir tür teknik üzerinde yoğunlaşmıştır. Google on yılı aşkın bir süre İMÇ'yi kullanmış ve 2016 yılının Kasım ayında nöral makine çevirisine (*Neural Machine Translation: NMT*) geçtiğini ilan etmiştir. Günümüzde Google, Yandex, Bing gibi endüstrinin büyük liderleri halen bu yöntemi kullanmaktadır.

### **2.5.3. Nöral Makine Çevirisi**

İfade tabanlı İMÇ'ler, Nöral Makine Çevirisinin (NMÇ) umut verici sonuçları görülene kadar en tercih edilen yöntemler olarak bilinmektedir (Koehn vd, 2007). MÇ için yapay sinir ağlarının kullanımı yıllar öncesinden başlamış (Castaño, 1997) olsa da çalışmalara, derin sinir ağları gibi büyük modellerin eğitilebilmesi mümkün olana kadar ara verilmiştir (Kalchbrenner, 2013, Cho vd., 2014b)

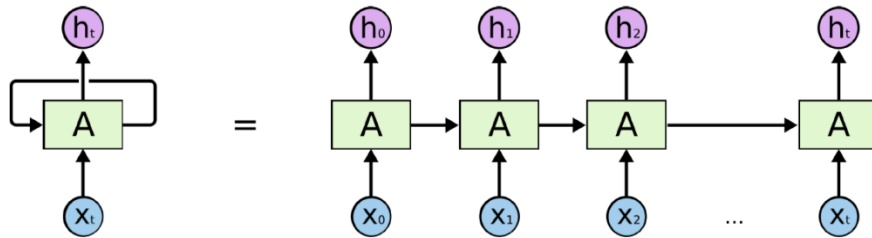
Beynin çalışma prensibini taklit eden derin öğrenme yöntemlerinin doğal dil işleme problemlerinde başarılı sonuçlar vermesiyle birlikte özellikle metin sınıflandırma ve duygu analizinde konvolüsyonel sinir ağları (*Convolutional Neural Network: CNN*) ve aşağıda bahsedilen kodlayıcı-kod çözücü gibi makine çeviri mimarilerinde ise özyineli sinir ağları (*Recursive Neural Network: RNN*) önerilmiştir (Young vd., 2018).

NMÇ, genel olarak bir şifreleme ve şifre çözme işi yapan **kodlayıcı-kod çözücü** (*encoder-decoder*) mimarisine dayanır. Kodlayıcı reel değerli bir rakamsal gösterimi hesaplamak için KD cümlesini okur ve kod çözücü ise önceden hesaplanmış bu gösterime göre her seferinde bir kelime olmak üzere HD çevirisini üretir (Şekil 2.18).



**Şekil 2.18: Kodlayıcı-kod çözücü modeli genel görünümü**

İlk model (Sutskever, Vinyals ve Le, 2014) HD cümleleri üretimi için KD cümlelerinin sabitlenmiş bir gösterimini kullanmıştır (Bahdanau, 2015). Bu mimarilerde sıralı bilgileri modellemek için **Özyineli Sinir Ağları** (ÖSA) kullanılmıştır. ÖSA'lar daha önceki adımda elde edilen çıktıların yeniden girdi olarak kullanılmasına imkan veren ve sahip oldukları bu döngü sayesinde cümle çevirilerinde daha önce geçmiş olan bir kelimenin hatırlanmasını mümkün kılan yapılardır. Bu bilgi bir sonra gelecek olan kelimenin tahmininde işe yaramaktadır. Aslında diğer sinir ağlarından çok fazla farkı olmayan ÖSA'lar için, kendisinin kopyalarından oluşan ve mesajı bir sonra gelene devreden yapılar oldukları söylenebilir. Şekil 2.19 ile özyineli sinir ağı yapısı ve döngüsünün açıldığındaki zincirli özelliği görsel hale getirilmiştir (Olah, 2015).



**Şekil 2.19: Bir ÖSA yapısı ve döngüsü açılmış haldeki görünümü**



Hochreiter ve Schmidhuber (1997) tarafından tanıtılmıştır. Zamanla geliştirilen bu ağlar oldukça yaygın bir şekilde kullanılmaktadır.

### ***Diziden-diziye (Sequence-to-sequence) modeller ve dikkat mekanizması:***

Diziden-diziye (*Seq2Seq*) yapısı makine çevirisi, metin özetleme ve resim yazısı oluşturma gibi işlerde çok sayıda başarı gösteren bir derin öğrenme modelidir. ‘Kodlayıcı-kodçözücü’ olarak da adlandırılan bu modeller, iki öncü metin olan Sutskever vd.’nin (2014) ve Cho vd.’nin (2014a) çalışmalarında açıklanmaktadır. Her iki çalışmada da **Uzun Kısa Vadeli Bellek: UKVB** ile kodlayıcı-kod çözücü ÖSA’lar kullanmış ve oldukça başarılı çeviriler elde edilmiştir. UKVB, standart ÖSA’lardan çok daha iyi çalışan bir özyineli sinir ağıdır. Örneğin “Almanya’da doğdum ve orada büyüdüm. En iyi konuştuğum dillerden biri ...” cümlesinin ‘Almancadır’ kelimesiyle bitmesi gerektiğini bir ÖSA tahmin edemez. Çünkü elde ettiği bilgiyle, o bilginin kullanılması gereken yer arasındaki mesafe oldukça uzamıştır. ÖSA’lar kelimenin bilgisini bir sonraki aşamaya aktarsa da o kelimenin cümle içinde ne kadar önemli olduğu bilgisini taşıyamazlar. UKVB’ler özellikle bu uzun mesafeli bağlantılar için tasarlanmışlardır. Zincire benzer yapıyla dört sinir ağı katmanına sahip UKVB, kendine has hesaplama yoluyla önemli olan bilginin tutulup önemsiz olanın atıldığı bir yöntemle çalışmaktadır. Ancak cümlelerin uzunluğu arttıkça UKVB’lerin verimliliği düşmektedir (Sutskever, Vinyals ve Le, 2014).

Diziden-diziye (*S2S*), bir dizi öğeyi (harfler, kelimeler vb.) alan ve başka bir öge dizisi olarak üreten bir modeldir. Şekil 2.18 ile, eğitilmiş bir modelin çalışma şekli görselleştirilmiştir. Görseldeki (Şekil 2.18) kodlayıcı, girdi dizisindeki her bir öğeyi ( $X_1, X_2, X_3, \dots, X_m$ ) işlemekte, yakaladığı bilgileri sayı dizilerinden oluşan bir vektöre dizmektedir. Tüm girdi dizisi işlendikten sonra bu vektör, tek tek ( $Y_1, Y_2, Y_3, \dots, Y_n$ ) üretilmek üzere kod çözücüye gönderilmektedir (Şekil 2.18). Bu modeller, örneğin altı kelimelik bir girdiyi dörk kelimelik farklı bir uzunluktaki çıktıyla eşleştirebilmek üzere tasarlanmıştır. Böyle bir çevirinin sıradan **Uzun Kısa Vadeli Bellek** (Hochreiter, 1997) mimarisiyle yapılabilmesi mümkün değildir. Bu nedenle diziden-diziye modelleri kullanılmaktadır. Sistemde kodlayıcı kelimeleri tek tek alarak

kelime yerleştirme (*word embedding*) algoritması kullanarak vektör haline dönüştürülür. Bu vektörlerde kelimeler anlambilimsel noktada bir temsil niteliği kazanır (Örn: ‘kral – adam + kadın = kraliçe’ vb.). Anlamsal mesafelere göre dizilen kelimelerin sayısal ifadeleri olan vektörler son safhada kod çözücü yardımıyla tek tek kelime haline getirilerek üretim aşaması gerçekleşmiş olur (Şekil 2.18).

Ancak bu tür sıralı hesaplamaların en büyük dezavantajı uzun cümleleri işlerken darboğazların (*bottleneck*) meydana gelmesidir. Yakın zamanlarda, yalnızca dikkat mekanizmalarına dayanan ve özyinelemeyi kullanmayan yeni bir model mimarisi olan Dönüştürücü (*Transformer*) önerilmiştir. Mimari üstün başarı göstermiş ve birkaç dil çiftinde dikkate değer kalitede sonuçlar verdiği kanıtlanmıştır (Vaswani vd., 2017).

#### ***Dönüştürücü (Transformer):***

Kodlayıcı-kodçözücü (*Encoder-Decoder*) mimarileri oldukça başarılı olmalarının yanında uzun cümlelerin çevirisinde sorunlar ortaya çıkmaktadır. Girdi cümlelerinin tek bir vektörle temsil ediliyor olması nedeniyle ilk geçen kelimeler bütün içinde önemini yitirmekte ve bu da istenilen çevirinin elde edilememesiyle sonuçlanmaktadır. Çünkü tüm kaynak sekansın doğru anlamlarıyla sabit bir vektöre sıkıştırılarak temsil edilebilmesi zordur. Bu mesele UKVB’lerle halledilme yoluna gidilse de istenilen düzeyde başarı elde edilememiştir.

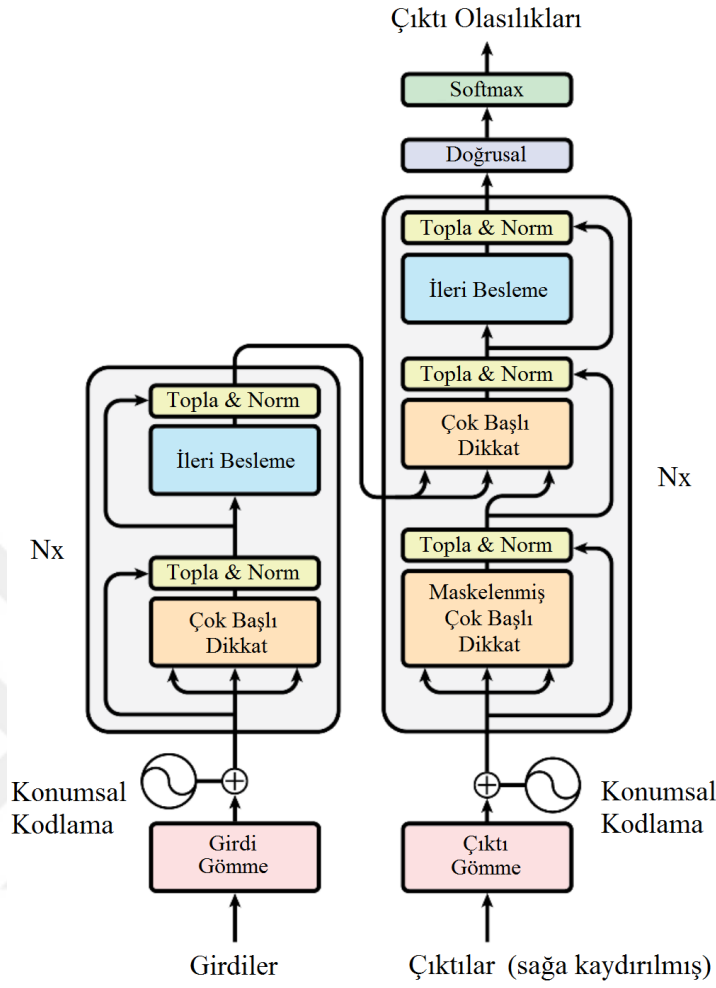
Bahdanau vd., (2014) nöral makine çevirilerinde ilk kez dikkat mekanizmasından bahsetmiştir. Dikkat mekanizmasındaki mantık bir değeri ararken kaynaktaki hangi noktalara daha fazla dikkat edilmesine karar vermektir. Bu ihtiyaca yönelik basit dikkat ağları mevcuttur. Örneğin; uzun kısa vadeli bellek problemleri üzerine yapılan bir çalışmada (Raffel ve Ellis, 2015) ‘basitleştirilmiş dikkat ağı’ terimi kullanılmıştır. Çift yönlü uzun kısa süreli bellek ağları üzerinde yapılan bir başka çalışmada ise (Zhou, vd., 2016) dikkat mekanizması (*Attention mechanism*) ifadesi yer almıştır.

Google, 2017 yılında ‘*Attention is All You Need*’ adlı bir makalesinde (Vaswani vd., 2017) ÖSA’ların yetersizliğine vurgu yapan yeni bir NMC mimarisi sundu. Makalede tanıtılan Transformer’ın farkı diğer transdüksiyon (*transduction*) yani bir

sekansı diğer sekansa çevirme işlemi yapan modellerin kalbi olan özyineli ÖSA'yı (*RNN*), UKVB'yi (*LSTM*), kapılı tekrarlayan hücre ağlarını (*GRU*) veya evrişimli sinir ağlarını (*CNN*) atıyor olması ve yerine yalnız dikkat (*attention*) mekanizmasını koyuyor olmasıdır. Bu yeni mimaride ÖSA'lara ihtiyaç duyulmaması sıralı işlem zorunluluğunu da ortadan kaldırmaktadır. Bu şekilde paralelizasyona izin veren mimari sayesinde sistemin eğitilmesi süreci oldukça hızlı hale gelmektedir.

Dönüştürücü, tüm girdi kelimeleri arasındaki ilişkiyi modelleyen ve bir kelimenin temsilini cümledeki tüm kelimelerle kıyaslayarak hesaplayan **öz-dikkat** (*self attention*) adlı bir mekanizma kullanır. Bu kıyaslama hesabı, diğer kelimelerin temsilinin mevcut kelimenin temsiline ne kadar katkıda bulunması gerektiğini belirleyen bir dikkat puanı üretir. Bu işlem sabit sayıda adım (katman) için tüm giriş kelimelerinde paralel olarak tekrarlanır. Kelimelerin hepsinin aynı anda işlenebiliyor olması sistemin uzun vadeli bağımlılıkları ÖSA'lardan çok daha iyi yakalamasına fırsat vermekte ve bu şekilde kelimeler arasındaki mesafenin önemi ortadan kalkmaktadır. ÖSA'ların dikkat mekanizması ile desteklendiği modellerde dahi ardışık çalışma prensiplerinden dolayı uzun girdili çevirilerde istenen başarı elde edilememiştir.

Çoğu seçkin nöral ardışık transdüksiyon modeli, kodlayıcı-kod çözücü mimarisine sahiptir (Bahdanau vd., 2014; Cho vd., 2014; Sutskever vd., 2014). Dönüştürücü mimarisinde kodlayıcı,  $(x_1, x_2, \dots, x_n)$  sembol temsilleri girdi dizisini ardışık gösterim dizisi olan  $z = (z_1, z_2, \dots, z_n)$  ile eşler. Kod çözücü,  $z$ 'yi girdi olarak alıp her seferinde bir eleman olacak şekilde  $y = (y_1, y_2, \dots, y_m)$  sembol çıktı dizisini oluşturur. Her adımda model, sonraki sembolü üretirken ilave girdi olarak önceki sembolleri tüketmektedir. Şekil 2.21 ile kodlayıcı (solda) ve kod çözücüde (sağda) öz-dikkat ekli, tam bağlantılı katmanlar kullanan dönüştürücü mimarisi (Vaswani vd., 2017) sunulmuştur.



**Şekil 2.21: Dönüştürücü modeli mimarisi**

**Kodlayıcı** (*Encoder*) altı adet aynı katmana ( $N=6$ ) sahiptir (Şekil 2.21) ve her katman iki alt-katmandan meydana gelir. Birincisi çok kafalı dikkat bloğu, ikincisi ise basit, pozisyon olarak tam bağlantılı ileri beslemeli yapay sinir ağıdır. Her iki alt-katmanın da etrafından bir artık (*residual*) bağlantı (He vd, 2016) uygulanır ve ardından katman normalizasyonu (Ba, Kiros ve Hinton, 2016) yapılır. Her alt-katmanın çıktısı  $\text{KatmanNorm}(x + \text{Altkatman}(x))$  şeklindedir ve  $\text{Altkatman}(x)$ , alt-katmanın kendisi tarafından uygulanan fonksiyondur. Bu artık bağlantıların işini kolaylaştırmak için modeldeki tüm alt-katmanlar ve gömme katmanları, boyutu  $d_{\text{model}} = 512$  olan çıktı üretir.

**Kod çözücü** (*Decoder*) de altı adet aynı katmana (N=6) sahiptir (Şekil 2.21). Her kodlayıcı katmanındaki iki alt-katmana ek olarak kod çözücü, kodlayıcı yığınının çıktısı üzerinde çok kafalı dikkat gerçekleştiren üçüncü bir alt-katman ekler. Kodlayıcıda olduğu gibi her iki alt-katmanın da etrafından bir artık bağlantı uygulanır ve ardından katman normalizasyonu yapılır. Sistemin sağladığı, ÖSA’larda bulunmayan, paralelizasyon imkanı ile tüm adımların aynı anda işlenebiliyor olması, bulunulan pozisyondan sonraki pozisyonlara da atlama ihtimalini doğurmaktadır. Bunu engellemek için kod çözücü yığınındaki öz-dikkat alt-katmanı da modifiye edilmiş ve bu maskeleye sonucunda, sonraki kelimelere geçilmeksizin, *i* pozisyonundaki tahminlerin yalnız *i* pozisyonundan önceki bilinen çıktılara dayanarak yapılması sağlanmıştır.

Öz-dikkat mekanizması doğrudan bu isimle geçmese de ilk olarak Cheng, Dong ve Lapata’nın (2016) makine okuması üzerine yaptıkları bir çalışmada (*Long Short-Term Memory-Networks for Machine Reading*) kullanılmıştır. Bu çalışmada odaklanılan kelimenin kendisinden önceki kelimelerle olan ilişkisini öğrenen bir model tasarlanmıştır. Makalede yer alan ve örnek metin üzerinde mevcut bir kelimenin önceki kelimelerle olan ilişkisini değerlendiren bir görsel Şekil 2.22 ile sunulmuştur.

The FBI is chasing a criminal on the run .  
The FBI is chasing a criminal on the run .  
The FBI is chasing a criminal on the run .  
The FBI is chasing a criminal on the run .  
The FBI is chasing a criminal on the run .  
The FBI is chasing a criminal on the run .  
The FBI is chasing a criminal on the run .  
The FBI is chasing a criminal on the run .  
The FBI is chasing a criminal on the run .

**Şekil 2.22: Kelimelerin önceki kelimelerle olan ilişkisi gösterimi**

Modelde “*The FBI is chasing a criminal on the run.*” cümlelerinde kırmızı ile gösterilen kelimeler (Şekil 2.22) o anda tespiti yapılmakta olanları; mavi renk ise koyuluğu nispetinde hafıza aktivasyonunu temsil etmektedir. Bu makine okuma sisteminde metin soldan sağa işlenmiş; hafıza ve dikkat için sığ muhakeme (*shallow reasoning*) uygulanmıştır. Ayrıca sistemin bir kodlayıcı-kod çözücü mimarisisiyle nasıl entegre edilebileceği tarif edilmiştir. Dönüştürücü (*Transformer*) mimarisinden (Vaswani vd., 2017) bir yıl önce sunulan bu model, dil modelleme, duygu analizi ve doğal dil çıkarımı üzerine yapılan testlerde en son yapılmış olan çalışmalara kıyasla denk; bazen de biraz üstü performans sergilemiştir.

#### **2.5.4. Derlem Tabanlı Makine Çevirilerinin Avantaj ve Dezavantajları**

Derlem tabanlı yöntemlerin avantajlı yanları maddeler halinde şu şekilde sıralanabilir:

- Derlem tabanlı bir sistemin geliştirilmesi kural tabanlı sistemlere kıyasla çok daha düşük maliyetlidir.
- Dilbilgisi noktasında uzman yardımına veya alan bilgisine ihtiyaç duymaz.
- İnsan emeği gerektiren kural veya sözlüklerin oluşturulmasına gerek yoktur. Sistemin eğitilmesi derlemler yardımıyla otomatik şekilde gerçekleşir. Bu da sistem kurulum maliyetlerini düşürmektedir.
- İMÇ sistemleri, diğer dil çiftleri veya mesleki alanlar için uyarlanabilir.
- Büyük dil modelleri sayesinde daha akıcı çeviriler üretebilir.
- Sistemin eğitiminde yeterli kaynak sağlanması durumunda farklı dil çiftleri için de kullanılması mümkündür.
- Eğitim verilerinde bulunan deyim veya dile has kullanımları yakalayabilir.
- Çeviri kalitesi, derleme daha fazla veri eklenerek daha iyi hale getirilebilir.

- Kelime anlamları seçiminde rolü olan n-gram modelleri, derlemlerin büyüklüğüne ve kalitesine bağlı olarak iyi sonuçlar verebilmektedir.

Derlem tabanlı modellerin dezavantajlı yanları ise maddeler halinde şu şekilde özetlenebilir:

- Kaliteli çeviriler için büyük ve kaliteli paralel derlemlere ihtiyaç vardır. Yeterince büyük olmayan derlemlerle oluşturulan çeviri ve dil modelleri, istenen çıktıları üretmez. Ayrıca birçok dil çifti için gerekli derlemler bulunamamaktadır.
- Eğitim verilerinin hangi alanla ilgili olduğu çeviri kalitesini etkilemektedir. Hangi tür çeviri yapılması gerekiyorsa o alana özgü verilere ihtiyaç vardır. Ancak, eklemek gerekir ki veri seyrekliği oldukça yaygın bir sorundur.
- Çekim ekleri ve yapım eklerine bolca rastlanan Türkçe gibi sondan eklemeli diller için mümkün olan en gelişmiş sistemlere ihtiyaç duyulmaktadır. Bu da donanım noktasında maliyetleri arttırmaktadır.
- Deyim veya dile özgü kullanımlar derlem tabanlı modellerde de zorluk çıkarmaktadır. Genel olarak çevirilerden istenen verimi alabilmek için büyük veriler üzerinde çözümlenebilecek yüksek maliyetli donanımlara ihtiyaç duyulur.

## **2.6. Çeviri Değerlendirmesinde Otomatik Yöntemler**

Çeviri kalitesinin ölçülmesinde en güvenilir yöntem insanlar tarafından yapılan çeviri değerlendirmeleridir. Bunun nedeni makine çevirisinin doğal dil davranışlarını yakalamada insan beynini taklit edebilecek düzeye ulaşamamış olmasıdır. Bu nedenle de makine çevirilerinde karşılaşılan en büyük güçlük olarak çeviri sistemlerinin başarı değerlendirilmesi ve çeviri kalitesinin objektif bir şekilde ölçülebilmesi olduğu söylenebilir. Çünkü doğal dillerin en büyük sorunu tam olarak matematiksel formüllerle ifade edilemiyor olmaları ve sosyal kabullere dayanan kullanımlar nedeniyle biçimsel ve anlamsal noktalarda tutarsızlıklar sergilemeleridir.

Ayrıca, bir cümlenin çoğu kez birden fazla şekilde doğru çevrilebiliyor olması ölçüm işlemlerinde üzerinde uzlaşmış yöntem bulunamayışını haklı çıkarmaktadır.

Makine çevirisinde kalite düzeyi sistem çıktılarının olması gerekene ne kadar yakın olduğuyula ilgilidir. Bir çevirmenin çıktığı kabul edilir düzeye getirmek için yapması gereken değişiklikler, değerlendirme sürecinde çeviri ölçütleri olarak kabul edilebilir. Ancak bu tür ölçümlerin subjektif oluşu, insan gücüne dayanması ve yüksek maliyetleri nedeniyle kullanılmayıp otomatik değerlendirme yöntemlerinin tercih edildiği görülmektedir.

Otomatik yöntemler içinde en yaygın olarak kullanılanlar BLEU, (Papineni vd., 2002) NIST, (Doddington, 2002) F Ölçütü, (Turian, Shen ve Melamed, 2003) METEOR, (Banerjee ve Lavie, 2005) ve WER'dir (Nießen vd., 2000). Bu yöntemler insanlar tarafından yapılan referans çeviriler ile sistemin ürettiği çeviriler üzerinde gerçekleştirilen farklı karşılaştırma ölçümlerine dayanır. Yukarıda sözü edilen yöntemlerden NIST, F Ölçütü; METEOR ise BLEU üzerinde değişiklik yapılarak geliştirilmiştir (Orhun, 2010).

### **2.6.1. BLEU / NIST**

Makine çevirilerinde insan değerlendirmesinin emek yoğun ve yüksek maliyetli olması nedeniyle Papineni vd. (2002) BLEU (*Bilingual Evaluation Understudy*) adını verdikleri bir yöntem geliştirmişlerdir. Yöntem, makine çıktıları ile çevirmenler tarafından çevrilen bir veya birden fazla referans metinlerin karşılaştırılması yoluyla çeviri kalitesini puanlandırmaktadır. Puanlandırma 0 ile 1 arasında olmakla birlikte literatürdeki bazı kaynaklarda BLEU skoru aralığı 0 ile 100 olarak da kullanılır. Puanın yüksekliği çeviri kalitesinin yüksekliğine işaret eder. Yöntem, çeviri başarısını değerlendirmede hatasız çözümler sunuyor olmasa da otomatik ve hızlı olması nedeniyle birçok doğal dil işleme görevinde yaygın olarak tercih edilmektedir.

BLEU yöntemi temelde kesinlik (*precision*) hesabına dayanır. Kesinlik, aday cümlede bulunan ve aynı zamanda referans cümle veya cümlelerde de yer alan toplam kelime (*unigram*) sayısının aday cümledeki toplam kelime sayısına

bölünmesiyle elde edilir. Ancak bu yolla yapılan hesaplama Şekil 2.23 ile örneklendirildiği gibi çeviri cümlesinde var olması gereken kelimeleri barındırdığı halde tamamen yanlış olan aday cümlelerin yüksek puan almasına neden olur.

Aday: bir bir bir bir bir bir bir
Referans 1: <u>Bir</u> işi <u>bir</u> kişi yapar.
Referans 2: Bir iş için tek kişi yeter.

**Şekil 2.23: 1-gram kesinlik örneği**

Örnekte (Şekil 2.23) yedi adet ‘bir’ kelimesine yedi kez referans cümlelerde rastlanmakta ve standart 1-gram kesinliği (*unigram precision*) 7/7 çıkmaktadır. Bu yanlış değerlendirmenin önüne geçmek için ölçümlerde aşağıdaki denklemde (2.19) görülen  $p_n$  değeri esas kabul edilir. Bu yolla doğru değerlendirme olan n-gram kesinliği 2/7 olarak belirlenmiş olur. Eşleşen n-gram sayıları, en fazla eşleştiği referanstaki sayıya göre belirlenir.

$$p_n = \frac{\sum_{C \in \text{Adaylar}} \sum_{Ngram \in C} \text{Adet}_{\text{bulunan}}(Ngram)}{\sum_{C' \in \text{Adaylar}} \sum_{Ngram' \in C'} \text{Adet}(Ngram')} \quad (2.19)$$

Doğal dillerde bir cümle bir başka doğal dile farklı biçim ve uzunluklarda çevrilebilir. Bu doğal durumun çeviri değerlendirmelerini olumsuz etkilememesi için ölçümlerde birden fazla referans cümle kullanılır. Fakat bazı aday cümlelerin referans cümlelerden daha kısa olduğu görülür. Bu durumun yanlış ölçümlere neden olmaması için ceza katsayısı yöntemi uygulanır. Aşağıdaki örnekte (Şekil 2.24) aday cümlelerin oldukça kısa olduğu görülmektedir. Adayın 2-gram kesinliği 2/2 ve 1-gram kesinliği ise 1/1 dir.

Aday: Hiç kimse.

Referans 1: Hiç kimse ona benzemez.

Referans 2: Onunla hiç kimse kıyaslanamaz.

### Şekil 2.24: 2-gram kesinlik örneği

Yukarıdaki gösterimde (Şekil 2.24) kesinlik hesabı, aday çevirideki ‘n-gram’lar ile referans çevirilerdeki ‘n-gram’ların kıyaslanması yoluyla yapılmıştır. N-gram cümlelerin kelime bazında n sayısı kadar alt parçalara bölünmesi işlemidir. Örneğin n = 1 olduğundaki 1-gram (*unigram*) hesabında metnin her kelimesi tek tek ayrı ayrı hesaplanmaktadır. Aynı şekilde n = 2 olduğunda da 2-gram (*bigram*) hesaplanırken metnin kelimeleri ikişer gruplar halinde parçalara ayrılır. (3 ve 4-gram için de aynı yöntem geçerlidir.) Örneğin; “Kediler balığı çok sever” cümlesinin ‘2-gram’ları “kediler, balığı”, “balığı, çok”, “çok, sever” şeklindedir. Aday cümledeki ‘n-gram’lar ile referans cümlelerdeki ‘n-gram’ların karşılaştırılması sırasında bulunan eşleşmeler bir daha hesaba katılmamak üzere yok sayılır.

Referans çevirilerden çok daha kısa bir aday çevirinin yüksek BLEU puanları almasını engellemek için uzunluk cezası (*brevity penalty*) adı verilen bir katsayı ile ağırlıklandırılması yapılmaktadır. Birden fazla referans cümle kullanıldığı için literatürdeki geri getirmenin (*recall*) uygulanamadığı BLEU yönteminde, kesinlik değerine göre hesaplama yapılır. Bunun için aşağıdaki formül (2.20) uygulanır. Formülde  $r$  referans cümlelerin,  $c$  ise aday cümlelerin uzunluğudur.

$$\text{uzunlukcezası} = \begin{cases} 1 & \text{eğer } c > r \\ e^{1-r/c} & \text{eğer } c \leq r \end{cases} \quad (2.20)$$

Kesinlik değeri ve ceza puanının tespitinden sonra aşağıdaki formül (2.21) takip edilerek BLEU puanı hesaplanır.

$$BLEU = uzunlukcezası \cdot \exp\left(\sum_{n=1}^N w_n \log P_n\right) \quad (2.21)$$

BLEU hesaplaması aday n-gramın referans n-gramlarla katlaştırılmasına dayanır. Türkçe gibi yapısal zenginliğe sahip eklemeli diller için bu yöntem yetersizdir. Bu nedenle BLEU yönteminin eksik yanlarını dikkate alınarak tasarlanmış daha derin analizler yapabilen uygulamalar geliştirilmiştir. Tantug, Oflazer ve El-Kahlout (2008) tarafından geliştirilen BLEU+ aracı ile, BLEU yöntemini Türkçe cümleleri işleyebilecek şekilde özelleştirilmiştir. Araçta aday ve referans metin kelimelerinin köklerine, çekim ve yapım eklerine bakılarak çözümlenmeler yapılabilmektedir. Ayrıca; kural tabanlı ve istatistiksel tabanlı çeviri modellerinin kıyaslanmasında BLEU puanının yeterli olmadığı ve bu yöntemle alınan veriler üzerinden yapılan iyileştirme çalışmaları sonrası insan değerlendirmelerinde, çeviri kalitesinde dikkate değer bir artış olmadığı gözlenmiştir (Callison-Burch ve Osborne, 2006). BLEU yöntemi çeviri kalitesinin ölçümünde yeterli olmasa da otomatik ve hızlı olması nedeniyle yaygın olarak kullanılmaktadır.

NIST metodu ise BLUE metodunun benzeridir. Eşleşen n-gramları sıklıklarına göre değerlendiren NIST yönteminde sık geçen 1-gram ile nadir geçen 1-gram aynı şekilde değerlendirilmez; nadir olanlar daha değerli kabul edilir. Ayrıca, BLEU hesaplanırken geometrik ortalama kullanılırken NIST, aritmetik ortalama kullanır (Doddington, 2002).

### 2.6.2. WER

Kelime Hata Oranı, (*Word Error Rate*) sözcük oranı konuşma tanıma ve makine çevirisinde yaygın kullanılan basit bir değerlendirme modelidir. Levenshtein mesafesinden (Levenshtein, 1965) esinlenen bu ölçütte sesbirimleri yerine kelimeler üzerinde değerlendirmeler yapılır. Levenshtein mesafesi, hedef kelimeyle kaynak kelime arasındaki benzerliği bulmak için karakter silme, ekleme ve değiştirme için gereken en az işlem oranıdır.

Office programlarının ve arama motorlarının yazım denetiminde kullanılan Levenshtein mesafesinin ‘Hanya’ ile ‘Konya’ kelimeleri arasında hesaplandığı düşünüldüğünde; (i) ‘H’ harfi ‘K’ harfi ile değiştirilerek ‘Hanya’ kelimesi ‘Kanya’ kelimesine; (ii) ‘a’ harfi ‘o’ harfi ile değiştirilerek de ‘Kanya’ kelimesi ‘Konya’ kelimesine dönüştürülmelidir. Dönüştürme süreci iki aşamada gerçekleştiği için sonuç 2 olmalıdır.

Farklı uygulamalar için farklı yöntemleri kullanılan WER, temelde şu denklem (2.22) ile hesaplanmaktadır:

$$WER = \frac{S + D + I}{N} \quad (2.22)$$

Denklemden, (2.22) ‘S’ değiştirilen karakter sayısını, ‘D’ silinen karakter sayısını, ‘I’ eklenen kelime sayısını ve ‘N’ ise referansta bulunan toplam kelime sayısını belirtir (Orhun, 2010).

İlk çeviri metriklerinden olan WER, MÇ çıktıları değerlendirilmede yeterli değildir. Çünkü birden fazla referanslı çevirilerden gelen bilgileri bir araya getiremez ve çeviriye eşlik eden kelime ve cümlelerin yeniden sıralanmasını doğru şekilde modelleyemez (Snover vd., 2009a).

### 2.6.3. METEOR

BLEU ölçütünün eksiklerini kapatmak için tasarlanan METEOR (*Metric for Evaluation of Translation with Explicit Ordering*: Açık Sıralı Çeviri Değerlendirme Metriği) makine çevirisi ile çevirmenler tarafından üretilen referans metinler arasında geliştirilmiş bir unigram eşleştirme yöntemini kullanır. METEOR, kelime-kelime eşleşmeleri kelime-kök eşleşmelerini ve eş anlamlı kelime eşleşmelerini dikkate alarak başlayan çok yönlü ve kademeli olarak artan kelime hizalama yöntemidir. Bu metrik daha gelişmiş eşleştirme stratejilerini kullanabilecek şekilde kolayca genişletilebilir. İki dizgi arasında tüm geliştirilmiş ‘unigram’ eşleşmeleri bulunduğu METEOR makine çevirisi içinde referans çeviriyle eşleşen

kelimelerin ne kadar doğru sıralı olduğunu direkt yakalamak için tasarlanan unigram-kesinlik, unigram-hatırlama ve parçalanma ölçüsü kombinasyonu yardımıyla, eşleşme için bir puanlama yapar. Kesinlik hesaplamaları BLEU'ya benzer. Hatırlama sözcük düzeyinde hesaplanırken kesinlik ve hatırlama puanlarını birleştirmede harmonik ortalamadan yararlanır. Ayrıca, METEOR daha uzun n-gram eşleşmelerini ödüllendirmekte; (Banerjee ve Lavie, 2005) diğer değerlendirme ölçütlerinden farklı olarak dilbilimsel süreçlerden de faydalanmaktadır (Elloumi vd., 2015).

## 2.7. İnsan Değerlendirmesi

Subjektif değerlendirme olarak da nitelendirilen insan değerlendirmesinde özellikle şu üç nokta dikkate alınır: (i) HD çevirisinin KD cümlesine göre anlamsal ve sözdizimsel doğruluğu; (ii) cümlenin anlamsal olarak doğruyken, sözdizimsel olarak yanlış olması; (iii) ve cümlenin hem anlamsal hem de sözdizimsel olarak yanlış olması (Bangalore, Bordel ve Riccardi, 2001). Değerlendirme sıralaması yaklaşımında akıcılık (*fluency*) ve doğruluk (*adequacy*) ölçümleri, gerekli değerlendirme açıklamalarıyla birlikte isteğe göre beş veya yedi katmanlı da olabilmektedir (Snover vd., 2009b) Cümlelerin akıcılık ve doğruluk değerleri, sözdizimsel ve anlamsal doğrulukla bağlı olarak değişmektedir. Bahsedilen kriterlere göre dizayn edilmiş dört katmanlı çevirmen derecelendirme ölçütü örneği, 'Çeviri Çıktıları ve Çevirmen Değerlendirme Ölçeği' (4.3) adlı başlık altında Tablo 4.3'te sunulmuştur.

Bu tez çalışmasında önerilen çeviri modeli Google, Bing ve Yandex çevirileri ile karşılaştırılmak suretiyle ölçümlendirilmektedir. Araştırmada, içinde eş anlamlı kelimeler bulunduran 40 adet isim grubu (*noun phrase*), 40 adet fiil grubu (*verb phrase*) içeren (80 adet) cümle girdileri (Tablo 4.1, Tablo 4.2) tez çalışmasında önerilen çeviri modeliyle İngilizceden Türkçeye çevrilmektedir. Cümlelerde en az bir, en fazla üç eş yazımlı bulunur. Aynı girdiler Google, Bing ve Yandex çeviriyle Türkçeye çevrilerek önceki listeyle birleştirilir (Ek: 1 – Ek: 10). Listede çıktıların hangi makine çevirisine ait olduğu değerlendirici olan A ve B çevirmenlerine bildirilmez. İki çevirmen tarafından Tablo 4.3 ile gösterilen derecelere göre

değerlendirilen çıktıların karşılık gelen ağırlık puanı, cümlelerin değerlendirme puanı olarak atanır. Ardından bu ağırlık puanlarının aritmetik ortalaması alınır. Böylece her makine çevirisi için 0 ile 100 aralığında belirlenmiş bir skor elde edilmiş olur.

## 2.8. İngilizce-Türkçe Makine Çevirisinde Karşılaşılan Güçlükler

Makine çevirisi (*machine translation*) terim olarak çeviri işleminin makine tarafından yapıldığını ifade ediyor olsa da aslında bir doğal dilin başka bir doğal dile doğru şekilde çevirisinin yapılabilmesi için her iki dilin dilbilimsel ve kültürel derinliklerine vakıf çevirmenlere ihtiyaç vardır. İkinci dünya savaşından bu yana yapılan tüm makine çevirisi çalışmaları bu gerçeğe işaret etmektedir. Çünkü bu güne kadar teknik imkânların elverdiği ölçüde yapılan tüm çalışmalar önceki çalışmaların eksik yanlarını kapatma motivasyonlarıyla gerçekleşmiştir.

Doğal diller sahip oldukları yapılar itibariyle birbirlerinden farklı özelliklere sahip olsalar da temelde benzer davranışlar sergiler. Örneğin; her doğal dilde mutlaka özne, nesne, fiil gibi benzer fonksiyonlar icra eden yapılar mevcuttur. Bu yapılar her dilde farklı yer, şekil veya anlam bütünlüğüyle temsil edilse de tutarlı davranışlarıyla farklı dillerdeki yapılarla da kıyaslanabilmeleri mümkün olmaktadır. Diller arasındaki bu davranış benzerliği için, makine çevirisinde iki ana yaklaşım olan kural ve derlem tabanlı yöntemlerin temel motivasyonu olduğu söylenebilir.

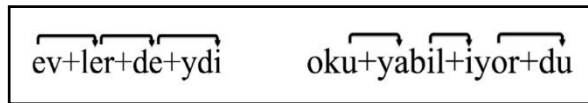
Dil davranışlarının bir makine tarafından değerlendirilebilmesi sürecinde genelde dört ayrı zorluktan bahsedilebilir. Bunlar biçimbilim (*morphology*), sözcükbilim (*lexicology*), sözdizim (*syntax*) ve ifade (*discourse*) aşamalarında gerçekleşmektedir.

Tüm bu aşamalar dört ayrı başlık halinde ele alınırken tez çalışmasında önerilmekte olan çeviri yaklaşımına uygun olarak özellikle İngilizce-Türkçe çeviri çalışmalarında karşılaşılabilecek zorluklara değinilmiştir.

### 2.8.1. Biçimbilim düzeyi

Herhangi bir dilden Türkçe gibi sondan eklemeli ve ünlü-ünsüz harflerin yan yana gelirken uyum içinde olması gereken bir dile yapılan çevirilerde karşılaşılan zorluklar kaynak dildeki İngilizce kelimelerle hedef dildeki Türkçe kelimeler ve eklerin eşleştirmesi aşamasında ortaya çıkmaktadır. Çünkü İngilizcede birkaç kelimedenden oluşan bir ifadenin veya cümlenin Türkçede tek bir kelimeyle ifade edilebilir olması çeviri sisteminin Türkçe kelime karşılıklarını çekim ve yapım eklerini ayırt ederek çözümlenebilecek şekilde tasarlanmış olmasını gerektirir.

Biçimbilim açısından Türkçe bitişken (*agglutinative*) bir dildir. Dilde biçimbirimler bir kök kelimeye tespih taneleri gibi eklenir. Türkçede önek bulunmamakta; ayrıca Almancadaki gibi bir dizi isim kökü birbiri ardına eklenerek yazılan birleşik isimlere de rastlanmamaktadır. Türkçede kelimeler yaklaşık otuz bin kadar kök kelimeye çok üretken bir şekilde bir dizi ek ekleyerek oluşturulur. Bir cümlede kullanıldıklarında kelimeler bir dizi yapım ve çekim eki alır. Örneğin; “yap+abil+ecek+se+k” kelimesi İngilizcede “*if we will be able to do (it)*” cümlesi ile ifade edilir. Hemen hemen tüm biçimbirimlerin kullanılan ünlüler ve biçimbirim sınırlarındaki ünsüzler yönünden farklı şekilleri bulunmaktadır. Örneğin; “paket+ten” kelimesinde ekin ilk ünsüzü ve ünlüsü birleştiği gövdenin son ünsüz ve ünlüsü ile uyum içinde olmak için ‘t’ ve ‘e’ olarak seçilmiş; “araba+dan” kelimesi ise bir ünlü harf ile bittiği için ekin ilk ünsüz harfi ‘d’ olarak kalmıştır. Ancak ünlü harf uyum için *a* olmak durumundadır. Ünlü uyumu Şekil 2.25 örnek cümlelerinde görüldüğü gibi soldan sağa silsile şeklinde gelişmektedir (Oflazer, 2016).



Şekil 2.25: Ünlü uyumunun silsile şeklinde çalışması örneği

Türkçe dilindeki sadece ünlü uyumu olmaksızın bile uygulanamayan yapım ve çekim eklerinin ortaya çıkardığı ifadelerin, İngilizce dilinde yalnız kelimeler ve kelimelerin cümle içinde buldukları yere göre karşılıklarının oluşturulabiliyor olması; bu dilleri çevirisi zor doğal dil çiftleri kategorisine sokmaktadır. Bu tür durumlarda insan çevirisine yakın çıktılar elde edebilmek için istatistiksel yöntem çevirilerde büyük ve kaliteli paralel derlemlere; kural tabanlı yöntemlerde ise kontrollü bir dil çalışması desteğine ihtiyaç vardır.

### **2.8.2. Sözcükbilim düzeyi**

Makine çevirilerinde sözcük düzeyinde en çok karşılaşılan sorun eş yazımlı kelimeler bulunduran cümlelerin çevirileridir. Çünkü kaynak dildeki eş yazımlı bir kelime hedef dilde birden çok kelimeye karşılık gelmektedir. Örneğin; İngilizcede ‘*crane*’ kelimesi hem ‘vinç’ hem de ‘turna’ anlamına gelmektedir. Ayrıca bu kelime fiil olarak kullanıldığında ‘boynunu uzatmak’ veya ‘vinçle kaldırmak’ anlamlarına da gelir. Bu kadar farklı şekillerde kullanılabilen ‘*crane*’ kelimesinin doğru çevirisi için bağlama duyarlı bir yaklaşım uygulanması gereklidir. Kelimeyi bağlamında değerlendiren makine çevirisi yöntemlerinin bu konudaki başarıları, özellikle son yıllardaki derin öğrenmeye dayanan nöral çeviri yöntemlerinde göze çarpıyor olsa da halen uzun mesafe bağımlılıkları veya uzak yapılarla sahip dillerde doğru kelime karşılıklarının bulunabilmesinde zorluklar yaşanmaktadır.

Ayrıca, dilin bağlı olduğu kültüre ait deyim veya ifadeler içinde geçen kelimeler de makine çevirilerinde sorun teşkil edebilmektedir. Hatta bazen kaynak dildeki bir kelimenin çevirisi hedef dilde bulunmayabilir. Örneğin Türkçedeki ‘mahremiyet’ kelimesinin, Japoncada karşılığı olan bir kelime bulunmamaktadır (Jurafsky, 2009).

### **2.8.3. Sözdizim düzeyi**

İngilizce cümlelerde öğeler ‘özne-yüklem-nesne’ şeklinde sıralanırken Türkçede bu sıra ‘özne-nesne-yüklem’ olarak gerçekleşir. Zaman, yer, vb. belirteç öğeler cümlelerin herhangi bir yerinde olabilir. Öğelerinin bu şekilde serbest olması, aynı

özelliğe sahip doğal dillerde olduğu gibi cümledeki isim öbeklerinin baş sözcüklerinin işleve göre hal ekleri almasıyla sağlanmaktadır.

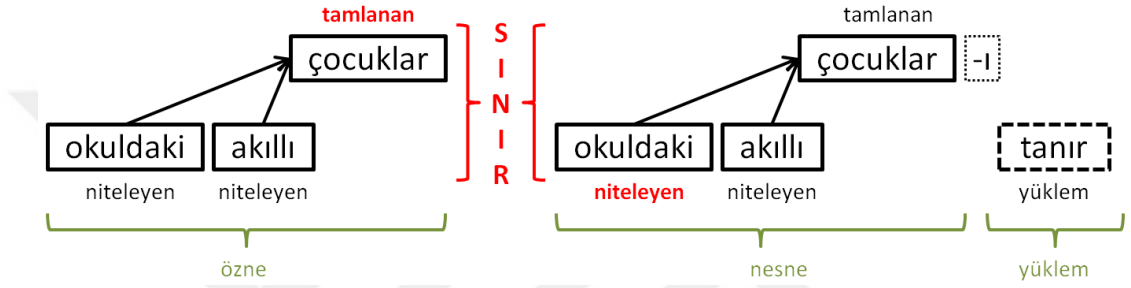
Şekil 2.26 ile öge sıralarının her biri için kabul gören çeviri beklentileri sunulmaktadır. Her cümlede gerçekleşen eylem Ekin'in Ayşe'yi görmesidir. Ögelerdeki sıra değişiklikleri, konuşma sırasında kabul gören çeviri beklentilerini kodlamaktadır (Oflazer, 2016).

- Ekin Ayşe'yi gördü.
- Ayşe'yi Ekin gördü. (Gören Ekin'di. Başka birisi değil!)
- Gördü Ekin Ayşe'yi. (Ama görmemesi gerekiyordu.)
- Gördü Ayşe'yi Ekin. (Zaten görmesini bekliyordum!)
- Ekin gördü Ayşe'yi. (Başkası da görebilirdi.)
- Ayşe'yi gördü Ekin. (Başkasını da görebilirdi!)

**Şekil 2.26: Çeviride öge sıralamalarına göre değişen anlamlar**

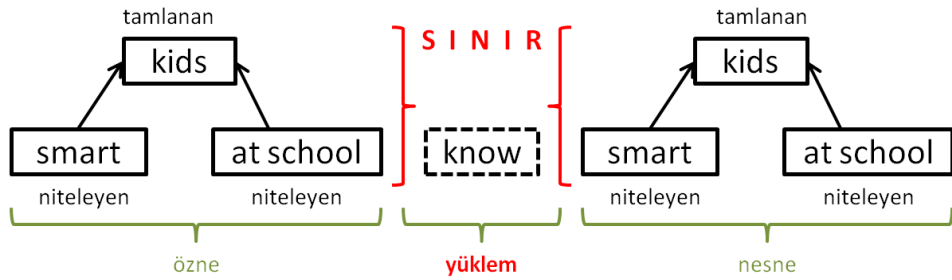
Türkçe dilinin öge sıralamalarındaki ÖNY (özne-nesne-yüklem) yapısı İngilizce dilinde farklı olarak ÖYN (özne-yüklem-nesne) şeklindedir. Bu durum her iki dildeki özne ve nesne yapılarının nitelenmeleri aşamalarına farklı etki eder. Yani Türkçede İngilizce gibi özne ile nesne arasında yüklem olmadığı için bu yapılar (özne ve nesne) oluşturulurken tamlanan kök kelime yalnız kendisinden önce gelen niteleme kelimeleriyle birleşerek anlam bütünlüğü oluşturur. Yani herhangi bir özne grubundaki kök kelimenin sadece soldaki kelimelerce tamlanıyor olması, aynı yolla kendisini yapılandıran nesne grubu ile karışmasını engeller (Şekil 2.27). İngilizcede ise yüklem ortada olduğundan kök kelimenin nitelenmesi hem baştan hem de sondan gerçekleşebilir. Yüklem ortada olması özne ve nesnelere niteleme kelimelerinin, kök kelimeleri her iki yönden tamlarken birbirleriyle karışmalarını engeller (Şekil 2.28).

Şekil 2.27 ile sunulan ‘Okuldaki akıllı çocuklar okuldaki akıllı çocukları tanır’ cümlesinde bulunan ‘okuldaki akıllı çocuklar’ ifadelerindeki ‘çocuklar’ kök kelimesi, önündeki ‘okuldaki’ ve ‘akıllı’ kelimeleri ile tamlanarak özne veya nesne fonksiyonunda anlam genişlemesine uğramıştır. Burada ‘çocuklar’ kök kelimesinin yalnız soldan niteleniyor olması, cümlede her iki isim grubu (özne ve nesne) için doğal bir sınır işlevi gören yapı meydana getirmektedir.



**Şekil 2.27: Kök kelimenin yalnız soldan nitelenmesiyle özne-nesne sınırının oluştuğu yapı**

Aşağıdaki İngilizce cümle örneğinde (Şekil 2.28) ise yüklem doğal sınır işlevi görmesi özne ve nesne gruplarının (*smart kids at school*) birbirlerinin sınırını ihlal etmeden, kök kelimelerinin (*kids*) her iki yönden de tamlanabilmesine izin vermektedir.



**Şekil 2.28: Yüklem ortada olmasıyla özne-nesne sınırının oluştuğu yapı**

Şekil 2.28 ile örneklendirilen ‘*Smart kids at school know smart kids at school.*’ cümlesindeki ‘*smart kids at school*’ isim grubundaki ‘*kids*’ kök kelimesi ‘*smart*’ ön nitelemesi (*premodifier*) ve ‘*at school*’ son nitelemesi (*postmodifier*) ile tamlanmıştır. Yüklemin her iki tarafında da oluşturulan bu isim grupları, konumlandıkları yere göre özne veya nesne özelliği kazanır.

Yüklemin özne ile nesne arasında olup olmaması, cümlenin diğer iki ana ögesi olan özne ve nesne gruplarının şekillenmesi sürecini de etkilemektedir. Bu durum dillerin yapı olarak uzaklaşmasındaki temel neden olarak gösterilebilir. Ayrıca tüm çeviri yaklaşımlarında, uzak yapılı dil çiftleri üzerine yapılan DDİ çalışmalarında istenen başarıların elde edilemediğinden daha önce bahsedilmiştir. Özellikle uzun cümleler söz konusu olduğunda çeviri sistemleri, farklı yapılandırılan bu dillerde tamamen yetersiz kalmakta; bu nedenle yapılacak makine çevirisi çalışmalarının sözdizim düzeyinde kapsamlı dil çalışmaları ile desteklenmesi gerekmektedir.

Özellikle Türkçe gibi sondan eklemeli ve morfolojik zenginliğe sahip diller için geliştirilen çeviri modellerinin, sınırlarının dışına çıkarak sözdizimsel ve anlambilimsel araçlar ile desteklenmesi gerekmektedir. Birbirinden uzak yapılı dillerin çevirisinde sistemlerin dil desteğine ne kadar ihtiyacının olduğu Tablo 2.3 ile gösterilen cümleler üzerinden örneklendirmek mümkündür. Tabloda farklı bir fiilin kullanılması ile nesne yapısının nasıl değiştiği görülmektedir.

**Tablo 2.3: Aynı kelime grubunun (nesne) üç ayrı hal eki ile temsili**

Örnek	Kaynak Dil	Hedef Dil
1	I like the team.	Ekibi seviyorum.
2	I have the team.	Ekibe sahibim.
3	I left the team.	Ekipten ayrıldım.

Verilen örneklerde (Tablo 2.3) görüldüğü gibi İngilizcelerde ‘the team’ isim grubu (*noun phrase*) herhangi bir ek almadığı gibi cümle sonunda konumlanmıştır. Ancak Türkçe çevirilerinin yapısı incelendiğinde cümlelerin fiiline göre; örneklerden birincisinin belirtme (-i hâli), ikincisinin yönelme (-e hâli) ve üçüncüsünün ise çıkma (-den hâli) durumlarını belirten hâl ekleri aldığı görülmektedir. Türkçe cümlelerdeki tüm nesnelere cümlenin başındadır. Ayrıca nesnelere, biçimbilim düzeyinde (2.8.1) bahsi geçen ünsüz uyumu durumu da geçerlidir (Tablo 2.3). Örneklerde gösterilen HD yapılarının fiile uygun şekilde doğru yerlerine koyulabilmesi için her iki dilin veritabanında da sözdizimsel bir ayrıştırmaya ihtiyacı vardır. Somers (1999) bu gibi durumlarda iyileştirme çalışmalarının zorluğundan bahsetmektedir.

Bu tez çalışmasında problem cümlesi olarak sunulmuş olan eş yazımlı kelimelerin bulunduğu cümle çevirilerinde ortaya çıkan zorlukların, halen yapılmakta olan tüm çeviri çalışmaları için geçerli olduğu söylenebilir. Tablo 2.4 ile verilen cümleler, aynı kelimenin üç farklı çevirisi örneği üzerinden İngilizce-Türkçe çeviri sistemlerinin dışarıdan dilbilimsel destekle güçlendirilmesi ihtiyacına işaret etmektedir.

**Tablo 2.4: Aynı kelimenin üç farklı çeviri örneği**

Örnek	Kaynak Dil	Hedef Dil
1	My friend <b>flies</b> to İstanbul.	Arkadaşım İstanbul’a <b>uçuyor</b> .
2	His pilot <b>flies</b> the plane.	Onun pilotu uçağı <b>uçuruyor</b> .
3	He doesn’t like <b>flies</b> .	O <b>sineklerden</b> hoşlanmaz.

Yukarıdaki (Tablo 2.4) KD cümlesindeki ‘*fly*’ kelimesi birinci örnekte ‘uçmak’, ikincisinde ‘uçurmak’ ve üçüncüsünde ise ‘sinek’ anlamlarıyla kullanılmıştır. Kelimenin hangi manada kullanıldığının bilinmesi Hem İngilizce hem de Türkçe diline ait sözdizimsel ve anlambilimsel analizlerle mümkündür. Bir başka benzer durum ise Türkçeye çevrildiğinde tamamen farklı yapım ekleri alarak isim, sıfat, zarf görevi alan kelimelerde görülmektedir. Tablo 2.5 ile örneklendirilmiş olan ‘dancing’ kelimesi birinci örnekte isim, ikinci örnekte sıfat ve üçüncü örnekte ise zarf görevindedir.

**Tablo 2.5: Tek bir yapının üç farklı çeviri örneği**

Örnek	Kaynak dil	Hedef dil	Yapı
1	The monkey likes <b>dancing</b> .	Maymun <b>dans etmeyi</b> sever.	<b>İsim</b>
2	The <b>dancing</b> monkey walked in.	<b>Dans eden</b> maymun içeri girdi.	<b>Sıfat</b>
3	The monkey walked in <b>dancing</b> .	Maymun <b>dans ederek</b> içeri girdi.	<b>Zarf</b>

Yukarıda (Tablo 2.5) ‘*dance*’ kökünden gelen ve ‘-*ing*’ eki olarak üç ayrı görevde kullanılabilen ‘*dancing*’ kelimesi, İngilizce form olarak hiçbir değişikliğe uğramazken Türkçe çevirisi olan yapıların farklı yapım eki aldıkları ve cümlede farklı pozisyonlarda buldukları görülmektedir.

Bahsedilen sorunların ortadan kaldırılması için geliştirilmiş olan yöntemlerden birkaç örnek vermek gerekirse; Kaji vd. (1992) belirsiz çeviri durumları oluşturan; örneğin ‘*play*’ gibi fiil veya isim olabilecek kelimeleri şablonlarda değişken halinde ve anlamsal kategoriler (spor, enstrüman vb.) altında tutmuştur. Richardson vd. (1993) kelime anlamındaki belirsizliği gidermek için örneklerin anlamsal bağlamlarını kullanmış; Sumita vd. (1993) KD dizgilerini veritabanı örnekleriyle eşleştirirken anlamsal mesafelerini ölçmek için bir eş anlamlılar sözlüğünden yararlanmış. Watanabe ise, (1993) KD girdisinin tam olarak yapılarına göre ayrıştırılmasını önermektedir ve ona göre bu prosedür daha fazla hesaplama karmaşıklığı pahasına da olsa cümlelerin sentezlenmesi problemini hafifletme potansiyeline sahiptir. (Hutchins, 2005).

#### **2.8.4. İfade düzeyi**

Deyim veya dile özgü kullanımlar tüm çeviri yaklaşımları için zordur. Genel olarak makine çevirilerinden insan çevirisine yakın çıktılar alabilmek için büyük veriler üzerinde çözümlene yapabilecek yüksek maliyetli donanımlara ihtiyaç vardır. Ancak bu zorluk yalnız deyimler vb. cümleler için geçerli değildir. Çevirilerde genel bir problem olan bu duruma ‘Kadın kedinin yanından geçerken havaya sıçradı.’ cümlesi örnek verilebilir. Cümle ‘*The woman jumped when she was walking by the cat.*’ (Kadın kedinin yanından geçerken zıpladı.) veya ‘*The cat jumped when the woman*

*was walking by.*’ (Kadın yanından geçerken kedi zıpladı.) şeklinde çevrilebilir. Bu tür ifade düzeyinde gerçekleşen anlam belirsizliklerine (*Dangling Ambiguity*) insan çevirisi dahil olmak üzere her tür çeviride rastlamak mümkündür.



## ÜÇÜNCÜ BÖLÜM

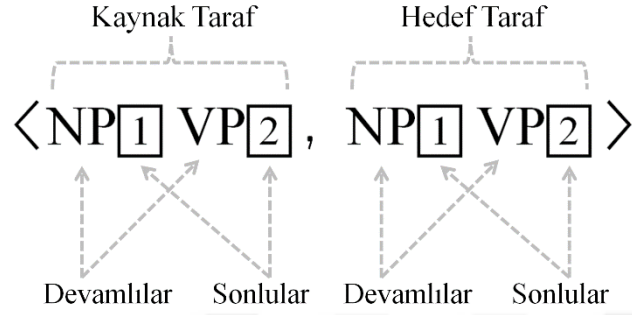
### İNGİLİZCE-TÜRKÇE KURAL TABANLI BİR ÇEVİRİ MODÜLÜ

Bu tez çalışmasında, temel fikir olarak istatistiksel makine çevirilerindeki (2.5.2) hiyerarşik ifade tabanlı istatistiksel makine çevirisine (Chiang, 2007) benzer bir model geliştirilmiştir. Ancak geliştirilen sisteminin eğitiminde Chiang'ın modelinin aksine paralel metinlerden yararlanılmayıp çevirmen tecrübelerine dayanan kural tabanlı yöntem tercih edilmiştir. Çalışmada önerilen Kural Matrisli Makine Çevirisi (KMMÇ) modelinde Chiang'ın çalışmasında da olduğu gibi senkronize içerikten bağımsız gramerden (*Synchronous Context-free grammar: SynCFG*) yararlanılmıştır.

Senkronize içerikten bağımsız gramer terimi yeni ve evrensel olmayan bir terimdir. Aslında 'syntax directed transduction grammars' (Lewis ve Stearns, 1968) veya 'syntax directed translation schemata' (Aho ve Ullman, 1969) olarak bilinmekte; ikinci isim muhtemelen en yaygın şekilde kullanılmaktadır (Chiang, 2006). Senkronize içerikten bağımsız gramer için, içerikten bağımsız gramerin (*Context-free grammar: CFG*) tek bir kelime dizisi yerine senkronize şekilde iki ayrı kelime dizisi üzerinden üretim yapabilen genelleştirilmiş hali olduğu söylenebilir. Bu nedenle iki farklı dil arasında yinelemeli ilişkiye dayanan birçok çalışma için, senkronize içerikten bağımsız gramerden yararlanılabilir. Aslında, 1960'ların sonlarında programlama dili derlemeleri için tasarlanmış (Chiang, 2005) olsa da sonraları makine çevirilerinde kullanılmıştır (Wu, 1997; Yamada ve Knight, 2001).

Şekil 3.1'de görüldüğü gibi senkronize içerikten bağımsız gramerde (*SynCFG*), hedef taraf ve kaynak taraf olmak üzere virgül ile ayrılmış iki bölüm vardır. Her iki bölümde de devamlılar (*nonterminals*) ve sonlular (*terminals*) mevcuttur. Kaynak taraftaki *NP* (*Noun Phrase* [isim ifadesi]) devamlısına (*nonterminal*) bağlı olan 1 numaralı sonlu (*terminal*) ile yine hedef taraftaki *NP* (*Noun Phrase*[isim ifadesi]) devamlısına bağlı olan ikinci 1 numaralı sonlu birbiri ile bağlantılıdır. Aynı şekilde kaynak taraftaki *VP* devamlısı da hedef taraftaki *VP* devamlısı ile irtibatlıdır ve bu

yolla bire bir eşlemesi gerçekleşen X sonlusu herhangi bir Y sonlusu ile değil; ancak bir başka X sonlusu ile ilişkilendirilmiş olur (Chiang, 2006).



**Şekil 3.1: SynCFG bileşenleri örneği**

Eş zamanlı eşleştirme yapabilen SynCFG için, İngilizceden Türkçeye çevrilmek üzere “*Cats like fish.*” cümlesi örnek gösterilebilir:

$$S \rightarrow \langle \text{NP}_1 \text{ VP}_2, \text{NP}_1 \text{ VP}_2 \rangle \quad (3.1)$$

$$\text{VP} \rightarrow \langle \text{V}_1 \text{ NP}_2, \text{NP}_2 \text{ V}_1 \rangle \quad (3.2)$$

$$\text{NP} \rightarrow \langle \text{cats, kediler} \rangle \quad (3.3)$$

$$\text{NP} \rightarrow \langle \text{fish, balık} \rangle \quad (3.4)$$

$$\text{V} \rightarrow \langle \text{likes, sever} \rangle \quad (3.5)$$

SynCFG’lerde de CFG’de olduğu gibi bir başlangıç sembolüne (*starting symbol*) ihtiyaç vardır. Üretim yapıldıkça devamlılar (*nonterminals*), yeni devamlılar ile değiştirilir. SynCFG’deki fark başlangıç sembolünün (S) iki tane oluşudur (3.6). (Kutulardaki 10 rakamı rastgele seçilmiştir.)

$$\langle S_{10}, S_{10} \rangle \quad (3.6)$$

İlk üretim (3.1) yapıldığında ‘S<sub>10</sub>’ yerine, içindeki devamlılar (NP<sub>1</sub> VP<sub>2</sub>) yazılır. Her üretim aşamasında kutu çiftlerinin sayıları da hiyerarşi gereği yeni sayılar ile değiştirilir (3.7).

$$\langle NP_{11} VP_{12}, NP_{11} VP_{12} \rangle \quad (3.7)$$

Bir sonraki üretim (3.2) gerçekleştiğinde ‘VP’ devamlısının yerine yeni devamlılar, onlara bağlı olan yeni numaralı kutularla birlikte yazılır (3.8).

$$\langle NP_{11} V_{13} NP_{14}, NP_{11} NP_{14} V_{13} \rangle \quad (3.8)$$

Bir sonraki üretim (3.3) uygulandığında devamlıları aynı olan NP<sub>11</sub> ile NP<sub>14</sub> uygulanmayıp kutulardaki sayılara göre sıra takip edilir (3.9) Yani NP<sub>11</sub> ile NP<sub>11</sub> üzerinden üretim gerçekleşir.

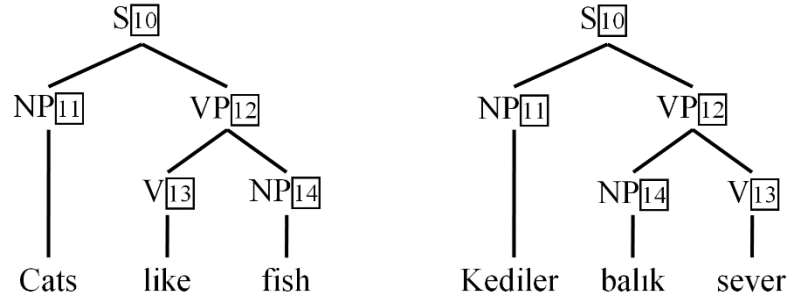
$$\langle \text{cats } V_{13} NP_{14}, \text{kediler } NP_{14} V_{13} \rangle \quad (3.9)$$

Diğer üretimler (3.4)(3.5) de sırasıyla gerçekleştikten sonra (3.10) kaynak cümle, hedef cümle olarak son halini almış olur (3.11).

$$\langle \text{cats } V_{13} \text{ fish}, \text{kediler } \text{balık } V_{13} \rangle \quad (3.10)$$

$$\langle \text{cats } \text{like} \text{ fish}, \text{kediler } \text{balık } \text{sever} \rangle \quad (3.11)$$

CFG’de olduğu gibi SynCFG’deki üretim aşamaları da çift ağaç gösterimi olarak şu şekilde görselleştirilebilir (Şekil 3.2):

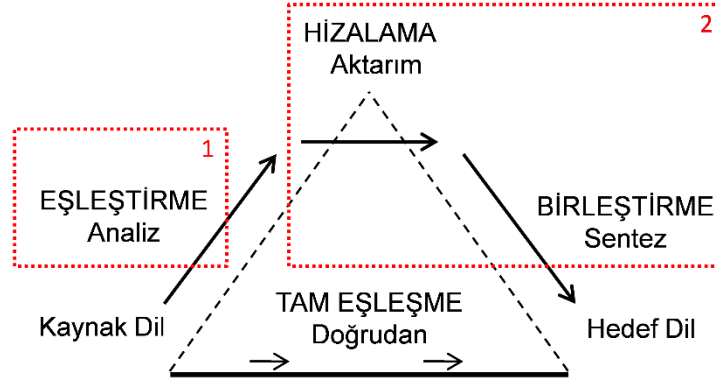


**Şekil 3.2: SynCFG’lerde üretim aşamaları ağaç gösterimi**

Bu tez çalışmasında önerilen KMMÇ modelinde yukarıda (Şekil 3.2) tarif edilen hiyerarşik ifade tabanlı yöntemle benzer bir analiz yöntemi uygulanmakta ve analiz aşamasında devamlıların bağlı olduğu kelimeler (sonlular) iki dilli sözlükten çekilerek üretim gerçekleştirilmektedir.

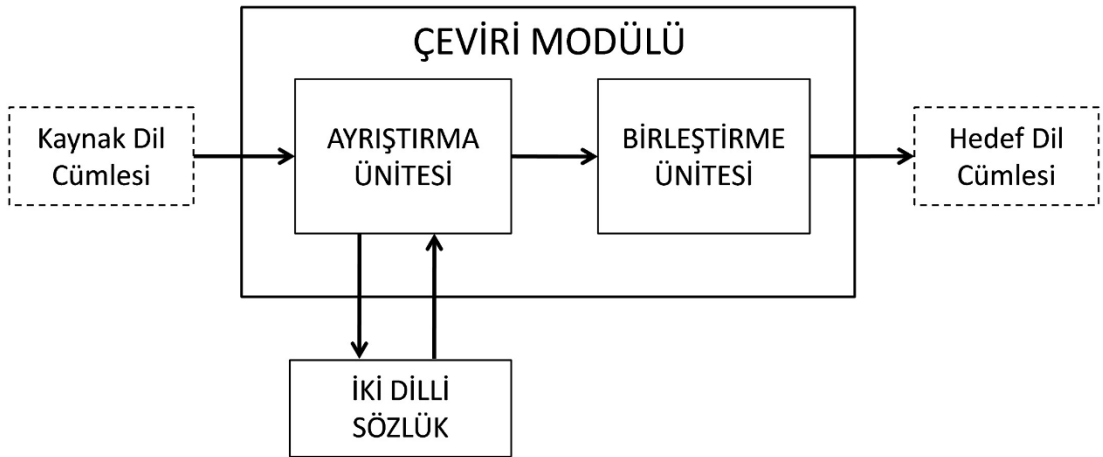
KMMÇ modelinde çeviri süreci iki aşamada yani iki ayrı üniteye gerçekleşir: (i) İngilizce kelimelerin Türkçeleri ile bu kelimeleri tutan devamlıların tespit edildiği **ayırıştırma ünitesi** ve (ii) kelimeleri temsil eden devamlıların belirli kurallara göre analiz edilerek üretimin sağlandığı **birleştirme ünitesi**dir (Şekil 3.4).

Sommers (1999) ÖTMC’nin aşamalarını (Eşleştirme, Hizalama, Birleştirme) Aktarım Makine Çevirisi aşamaları (Analiz, Aktarım, Sentez) üzerinden Vauquois üçgeni ile tarif etmiştir. Üçgende ÖTMC modeli aşamaları, aktarım yöntemindeki sınıflandırmaların hemen üzerinde büyük harflerle yazılmıştır (Şekil 3.3). Çalışma prensipleri farklı olsa da önerilen KMMÇ modelini yine benzer şekilde tarif etmek gerekirse, Eşleştirme ve Analiz süreci için birinci aşama (Ayrıştırma Ünitesi [1]); Hizalama ve Aktarım ile Birleştirme ve Sentez süreci için ise ikinci aşama (Birleştirme Ünitesi [2]) olduğu söylenebilir (Şekil 3.3).



**Şekil 3.3: KMMÇ modelinin Vauquois üçgeni üzerinden aşamaları**

Bu haliyle KMMÇ modeli, **kodlayıcı** ve **kod çözücü** (*encoder-decoder*) mimarisine sahip Nöral Makine Çevirisine (2.5.3) benzer bir ikili yapıya sahiptir. NMÇ, genel olarak bir şifreleme ve şifre çözme işi yapan **kodlayıcı** ve **kod çözücü** (*encoder - decoder*) mimarisine dayanır. Kodlayıcı reel değerli bir rakamsal gösterimi hesaplamak için KD cümlesini okur ve kod çözücü ise önceden hesaplanmış bu gösterime göre her seferinde bir kelime olmak üzere HD çevirisini üretir. Ancak önerilmekte olan KMMÇ sistemi, NMÇ'nin aksine derlem tabanlı değildir ve iki dilli sözlüğe ihtiyaç duyar. KMMÇ modelinin ayrıştırma ünitesi, birleştirme ünitesi ve iki dilli sözlüğü içeren genel görünümü Şekil 3.4 ile sunulmuştur.



**Şekil 3.4: KMMÇ modelinin genel görünümü**

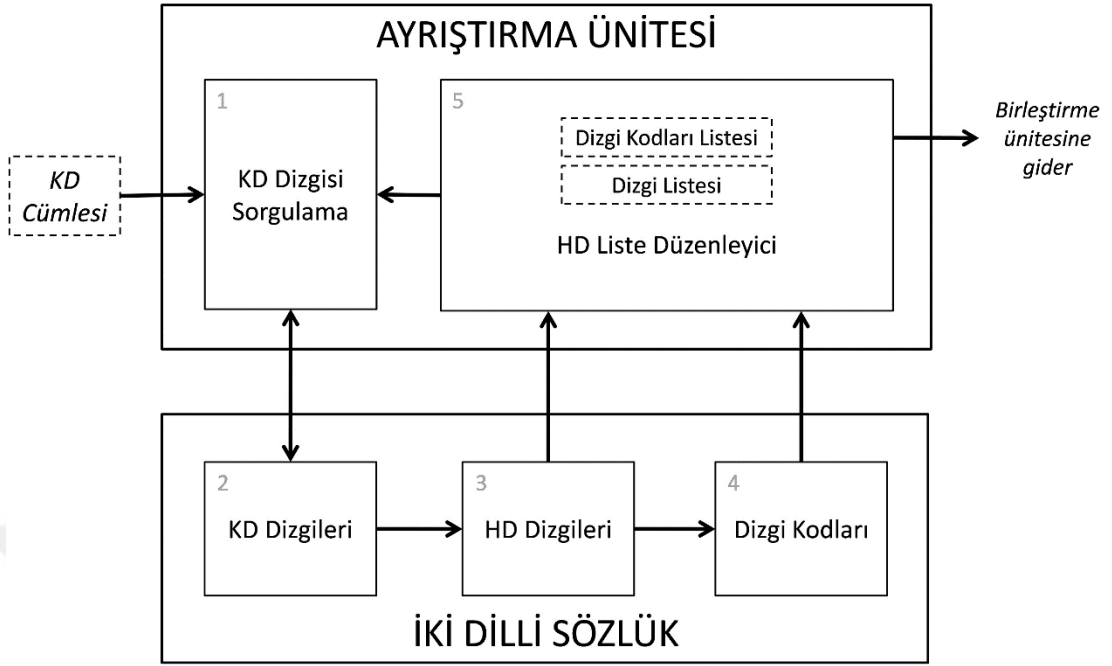
Çeviri modülüne (Şekil 3.4) girdisi yapılan KD cümlesi, ayrıştırma ünitesine gelir. Ayrıştırma ünitesi iki dilli sözlük yardımıyla KD girdisini gerekli ek, kelime ve ifade dizgilerine (*strings*) ayırmak ve bu dizgileri kodlar (*nonterminals* [devamlılar]) ile eşleştirmekten sorumludur. Eşleştirilmesi yapılan kodlar listesi ve Türkçe dizgiler listesi bir sonraki aşama olan birleştirme ünitesine gönderilir.

Birleştirme ünitesi, dizgileri kodları üzerinden analiz eder ve bu dizgilerin hangi sıra ile birleşeceklerine ve hangi yapım veya çekim eki alacaklarına karar verir. Böylece ayrıştırma ünitesinde kodlar verilerek parçalara ayrılan KD dizgisi, birleştirme ünitesinde bu kodlar yardımıyla bir araya getirilerek HD dizgisinin üretimi gerçekleştirilmiş olur.

Bu tez çalışmasında önerilmekte olan çeviri modülü ‘Ayrıştırma Ünitesi’ (3.1) ve ‘Birleştirme Ünitesi’ (3.2) olarak iki başlık altında sunulmuştur. Ayrıştırma ünitesi konusu (i) ‘KD dizgisi sorgulama birimi’ (3.1.1), (ii) ‘HD liste düzenleyici birimi’ (3.1.2) ve (iii) ‘iki dilli sözlük’ (3.1.3) alt başlıklarıyla ele alınmış; birleştirme ünitesinin alt başlıkları ise (i) ‘Çoklu kod sorgulama birimi’ (3.2.1), (ii) ‘Birleştirme matrisi’ (3.2.2), (iii) ‘Kod ve dizgi üretim’ (3.2.3), (iv) ‘Yapım-çekim eki ilave’ (3.2.4), (v) ‘Birleştirme ünitesinde HD üretim aşamaları’ (3.2.5) ve (vi) ‘Üretim aşamasında eş yazımlı kelimeler’ (3.2.6) olarak belirlenmiştir.

### **3.1 Ayrıştırma Ünitesi**

Ayrıştırma ünitesi KD girdisini gerekli ek, kelime ve ifade dizgilerine ayırmak ve bu dizgileri kodlar (*nonterminals*) ile eşleştirmek için üç ayrı birimden yararlanır. Bunlar KD dizgisi sorgulama birimi, iki dilli sözlük ve liste düzenleyici birimdir. Şekil 3.5 ile ayrıştırma ünitesinin bileşenleri ve aralarındaki ilişkiler görsel hale getirilmiştir.



**Şekil 3.5: Ayırıştırma ünitesi ve bileşenleri**

Ayırıştırma ünitesindeki (Şekil 3.5) KD dizgisi sorgulama birimi [1] KD cümlesini alır ve bu dizgideki ifade, kelime veya ek dizgilerini iki dilli sözlükte arar. Bulunan KD dizgilerinin (ifade, kelime, ek) [2] sözlükteki karşılığı olan HD dizgileri [3] ve dizgi kodları [4] HD liste düzenleyiciye [5] gönderilir. HD liste düzenleyici gelen HD dizgilerini dizgi listesinde, dizgi kodlarını da dizgi kodları listesinde sıralar ve bu iki listeyi çeviri modülünün ikinci ve son aşaması olan birleştirme ünitesine gönderir.

### 3.1.1. KD dizgisi sorgulama birimi

Ayırıştırma ünitesi (Şekil 3.5) sınıfı içinde tanımlanmış bir dizgi bulma metodu, KD dizgisindeki ifade, kelime ve ekleri iki dilli sözlükte aratılabilecek sorgular haline getirir. Metod, KD girdisini sözlükte tararken mümkün olan en az sayıda sorgu döndürecek şekilde çalışır. Bunun için, kelime arası boşluk karakterler de dâhil olmak üzere aramasını tüm girdi cümlesi karakterlerinden başlayıp her seferinde dizgi başından bir karakter eksilterek yapar. Sözlükteki dizgiyle eşleşen parçayı girdi dizgisinden çıkardıktan sonra kalan dizgiyi tekrar sorgularken yine tüm girdi karakterlerinden başlayıp her seferinde dizgi başından bir karakter eksilter. KD

dizgisi sorgulama birimi [1] bu döngüyü KD dizgisinde liste düzenleyiciye [5] gönderilecek dizgi kalmayana kadar; bir başka deyişle KD dizgisinde sıfır karakter kalana kadar devam eder (Şekil 3.5).

Her seferinde tüm girdi karakterlerinden başlanarak dizginin sözlük taramasından geçirilmesi, analiz edilen KD cümlesinde birden çok kelime barındıran dizgilerin öncelikli olarak tercih edilmesini sağlar. Böylece KD dizgisi sorgulama birimi, KD dizgisinde ‘*crane*’ ve ‘*fly*’ kelimelerini yan yana bulduğunda ‘dozer’ ve ‘sinek’ anlamına gelen iki ayrı dizgi olarak görmeyip ‘tipula sineği’ anlamına gelen ‘*crane fly*’ ifadesini tek dizgide yakalamış olur. Bu metod yalnız birden fazla kelime içeren deyim veya kültüre ait bazı kullanımların gereksiz yere dizgilere parçalanarak yanlış analizlere yol açmasını engellemekle kalmayıp tek bir kelime içerisinde oluşabilecek yanlış çözümleme durumlarını da ortadan kaldırır. Örneğin yaratmak anlamına gelen ‘create’ kelimesinin sonundaki üç harf (ate) sözlükte bulunduğu için KD dizgisinden çıkarılması durumunda kelimedен geriye kalan ‘cre’ dizgisiyle hiçbir sözlük eşleşmesi yapılamayacak ve bu da daha en baştan girdi dizgisinin doğru analizinin yapılabilmesini imkânsız hale getirecektir.

### 3.1.2. HD liste düzenleyici birimi

Ayrıştırma ünitesi sınıfı içinde tanımlanmış bir başka metod olan liste düzenleyici, iki dilli sözlükten elde edilen HD dizgileri [3] ve dizgi kodlarını [4] liste düzenleyicideki [5] listelere (dizgi listesi, dizgi kodları listesi) tüm elemanları sağdan sola olacak şekilde sıralar (Şekil 3.5). Kod ve dizgilerin tersten sıralanıyor olması KD dizgisi sorgulama biriminin, KD cümlelerini sözlükte sorgularken bulunduğu dizgileri cümlenin sonundan başlayarak çıkarmayı tercih etmesinden kaynaklanır. Çünkü çeviri modülünün ikinci ve son aşaması olan birleştirme ünitesinde de cümleler sondan başlayarak analiz edilmektedir.

Söz konusu listeler içinde İngilizce dizgi listesi de (KD cümlesi dizgilere ayrılırken) oluşturulmuş olsa da modülünün herhangi bir aşamasında kullanılmaz. Aşağıdaki (3.12) “*more mischievous cat*” örneğinde görüldüğü gibi kod ve dizgiler, köşeli parantezlerle oluşturulan (python) liste veri tipinde tutulur. Bu sayede her veri

tipinden eleman saklanabilir ve bu elemanlar üzerinde yapılacak deęişiklikler daha kolay hale gelir.

```
[[ 'cat',      'mischievous', 'more',  '|' ],      (3.12)
 [ 'Na',      'JJ',          'JJ'    '|' ],
 [ 'kedi',    'yaramaz',    'daha',  '|' ]]
```

Verilen iç içe liste (3.12) bir ayrıştırma ünitesi çıktısıdır (İngilizce liste burada eklenmiştir). HD liste düzenleyici [5] (Şekil 3.5) HD dizgilerinin [3] ve dizgi kodlarının [4] listelenmesi işlemini yine bir liste (3.12) içinde gerçekleştirir. Bu dış liste sayesinde listeler bir arada tutulur.

Örnekteki (3.12) “*more mischievous cat*” dizgisi üç parçaya bölüdüğü halde iç listeler dört elemanlıdır. Daha sonraki örneklerde de karşılaşılabilecek olan bu kod ‘|’ liste düzenleyici tarafından her listenin başına eklenir. Bu yolla eşleştirme ünitesinden çıkan listeler en az iki eleman tutmuş olur. Bunun sebebi sonraki aşama olan birleştirme ünitesi kod analizlerinde en az iki elemana ihtiyaç duyulmasıdır.

Girdiler taranırken sıklıkla eş yazımlı (birden fazla anlamı olan) kelimelere rastlanır. HD liste düzenleyici, kelimelerin diğer anlamlarını da kodlarıyla birlikte sözlükten çıkarıp kod listesine ekler. Mesela ‘*crane*’ kelimesi kod listesine ‘N’, ‘Na’ ve ‘V2’; Türkçe dizgi listesine ise ‘dozer’, ‘turna’ ve ‘boynunu uzat’ olarak işlenir. Eş yazımlı bir kelime en fazla beş ayrı çevirisi ile listelere eklenir. Aşağıda “*my old car’s doors*” ifadesinin ayrıştırma ünitesi çıktısı verilmiştir (3.13). Listede ‘old’ kelimesi iki farklı çevirisi ile listelenmiştir. Ancak bu çevirilerden hangisinin tercih edileceği ikinci aşama olan birleştirme ünitesine bırakılır.

```
[[ 's',      's',      'door',  's',      's',      'car',    'old',    'old',    'my',    '|' ],      (3.13)
 [ 'PLR',    'TPS',    'N',     'S_',     'BE',     'N',     'J',     'Ja',     'DT',    '|' ],
 [ '_ler',   ",        'kapı',  '_in',    '_dir',   'araba', 'eski',   'yaşlı', '-m',    '|' ]]
```

Ayrıştırma ünitesi iki dilli sözlükte sorgulanan kelime veya ifadelerde bulunan yapım veya çekim eklerini de kodlar. Yani yukarıdaki (3.13) “*my old car’s doors*” (eski arabamın kapıları) ifadesinde geçen “*car’s*” kelimesindeki sahiplik eki “-’s” ile “*doors*” kelimesini çoğul yapan ek “-s”, hem kod hem de Türkçe dizgi listesinde yerini alır. Böylece sözlükte, kelime sadece “*door*” olarak kaydedilmiş olsa da “*doors*” kelimesinden bağımsız bir kelime olarak görülmesinin önüne geçilmiş olur. Aynı durum “*car*”, “*car’s*” vb. kelimeler için de geçerlidir. Ayrıca, “*my old car’s doors*” ifadesindeki eklerin (-s, -’s) sadece tek bir kodla temsil edilmedikleri görülmektedir (3.13). Çünkü “*car’s*” ifadesindeki “-’s”, hem ‘olmak’ anlamına gelen İngilizcedeki ‘*is*’ kelimesinin kısaltılmışı “-’s” hem de sahiplik bildiren “-’s” olabilir. Mezkur ek, “*car’s door*” ifadesinde sahiplik bildirirken; “*Car’s here.*” (Ev buradadır.) cümlesinde, olmak anlamına gelen, ek fiil görevindedir. Görüldüğü gibi ekin, bu iki yapıdan hangisi olduğu, kendisinden sonra gelen kelime veya ifadelere bağlıdır. Ancak bu iki yapıdan hangisinin kullanılacağına ayrıştırma ünitesinde karar verilmez. Bu nedenle sözdizim ayrıştırıcı her oluşan ihtimali de kod ve Türkçe dizgi listelerine ekleyerek, analiz edilmek üzere birleştirme ünitesine gönderir.

### 3.1.3. İki dilli sözlük

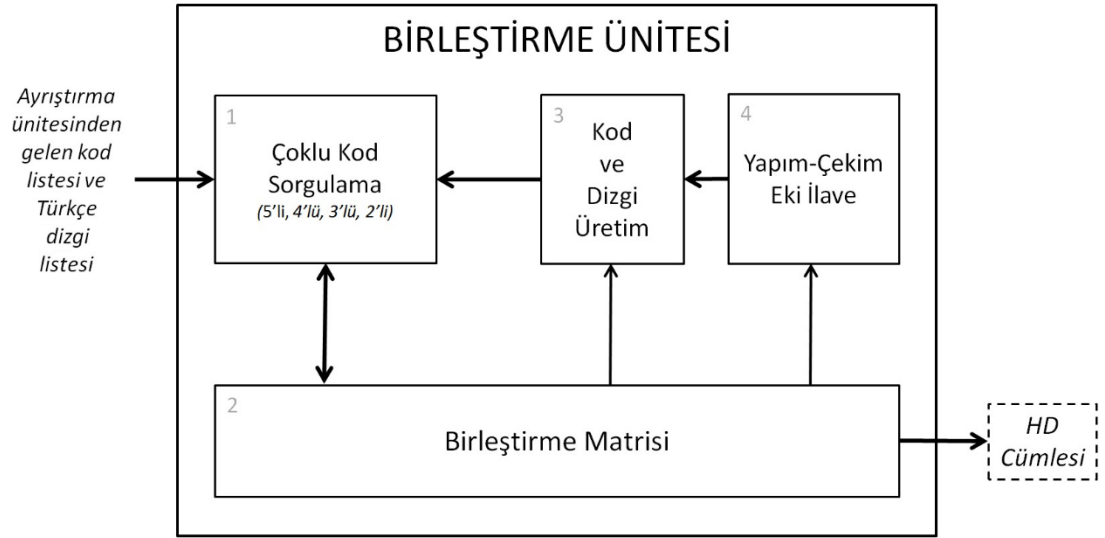
SQLite veritabanı olan iki dilli sözlük ‘kod’, ‘ingilizce’ ve ‘turkce’ sütunlarını içeren bir tablodur. KD dizgisi sorgulama birimindeki dizgi sorgulama metodu, kelime sorgularında SQL’den yararlanır. Veri yönetiminin en hızlı şekilde yapılabilmesini sağlayan bu dil sayesinde sürekli geliştirilme ihtiyacı olan sözlükteki dizgilerin ekleme veya çıkarma işlemleri kolayca yapılabilir.

Sözlükte her İngilizce dizginin bir kodu ve bir çevirisi vardır. İki veya daha fazla sayıda aynı dizginin sözlükte farklı kodlar aldığı durumlar yalnız eş yazımlı (‘*crane*’, ‘*fly*’ vb.) kelimelerin bulunduğu ifadelerde veya benzer karakterler içeren gramer yapılarında görülür. Örneğin “*doors*” kelimesindeki “-s” eki, sözlükte karşılık olarak hem geniş zamanda üçüncü şahıslar için fiile eklenen ‘TPS’ (*Third Person Singular*) koduyla; hem de çoğul eki olan ‘PLR’ (*Plural*) kodu ile eşleştirilir (3.13).

Sözlükteki İngilizce kelimeler için uygun kod ve çeviriler seçilirken biçimbilim, sözcükbilim, sözdizim ve sosyal kullanım düzeyinde değerlendirmelerden yararlanır. Bu sebeple iki dilli sözlük içeriği oluşturma veya geliştirme çalışmaları, alanında uzman ve her iki dilin de kültürüne hâkim kişilerce yapılmalıdır.

### 3.2. Birleştirme Ünitesi

Çeviri modülünün ikinci ve son aşaması olan birleştirme ünitesi, ayrıştırma ünitesinden gelen kod ve dizgi listelerinden HD çıktısını üretebilmek için dört ayrı birimden yararlanır. Bunlar (i) çoklu kod sorgulama, (ii) birleştirme matrisi, (iii) kod ve dizgi üretim ve (iv) yapım-çekim eki ilave birimleridir (Şekil 3.6).



Şekil 3.6: Birleştirme ünitesi genel görünümü

Birleştirme ünitesinde (Şekil 3.6) bulunan çoklu kod sorgulama birimi [1] ayrıştırma ünitesinden gelen kod listesi ve Türkçe dizgiler listesini birleştirme matrisinde [2] analiz eder. Birleştirme matrisi, değişiklik yapılması gereken dizgileri kod ve dizgi üretim birimine [3] ve ilave edilmesi gereken yapım veya çekim eklerini de yapım çekim eki ilave birimine [4] yönlendirir.

Birleştirme ünitesi, sahip olduğu birimler arası ilişkinin doğru tarif edilebilmesi için ‘Çoklu kod sorgulama birimi’, ‘Birleştirme matrisi’, ‘Kod ve dizgi üretim’ ve ‘Yapım-çekim eki ilave’ başlıklarıyla ele alınmıştır.

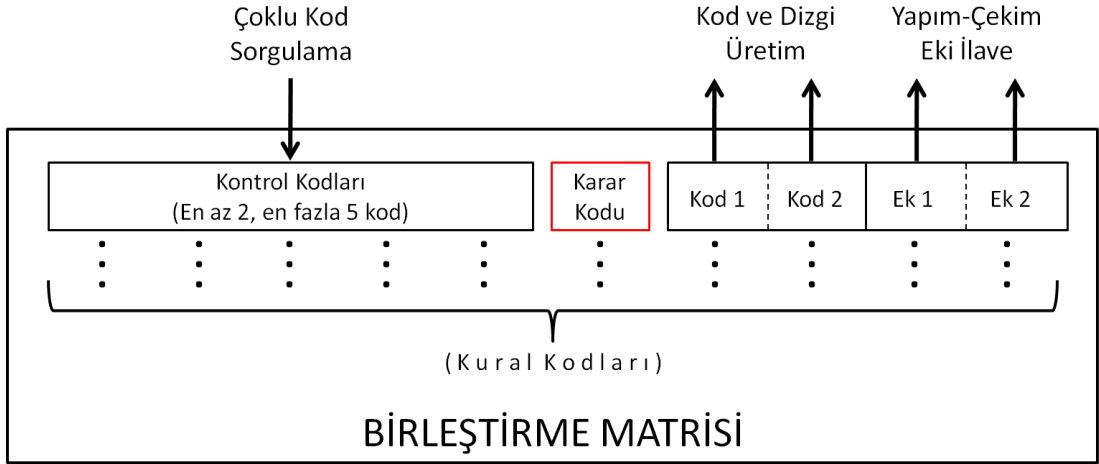
### 3.2.1. Çoklu kod sorgulama birimi

Ayrıştırma ünitesinden çıkan kod listesi ve Türkçe dizgi listesi ilk olarak çoklu kod sorgulama birimine (Şekil 3.6) gelir. Birimde tanımlanmış olan çoklu kod sorgulama metodu listelerdeki hangi kodların ve dolayısıyla hangi dizgilerin önce birleşeceğini bulmak için birleştirme matrisi içerisinde kod taraması yapar. Kod taraması, kod listesi elemanlarının (2’li, 3’lü, 4’lü veya 5’li gruplar halinde) birleştirme matrisi içindeki kontrol kodlarıyla (Şekil 3.7) eşleşmelerinin aranması işlemidir. Eşleşen kodlar bulunduktan sonra sıra, **birleştirme matrisinin** kod ve dizgi listeleri üzerinde yapacağı analize gelir.

### 3.2.2. Birleştirme matrisi

Birleştirme matrisi (Şekil 3.6) listeler üzerinde dört aşamalı bir analiz gerçekleştirerek (i) hangi kodlar ile bu kodlara bağlı dizgiler üzerinde değişiklik (birleştirme, değiştirme, silme vb.) yapılacağına; (ii) birleşecek veya değişecek kodların yerine hangi kodların geleceğine, (iii) dizgilerin hangi yapım-çekim eki alacağına (iv) ve yapım-çekim ekleri almış dizgilerin hangi sıra ile birleşeceğine karar verir.

Bu dört aşamalı analizin nasıl gerçekleştiğini tarif edebilmek için birleştirme matrisinin, birleştirme ünitesi içinde tanımlanmış bir liste veri yapısı olduğundan bahsetmek gerekir. Bu yapı, elemanları binlerce liste olan bir iç içe listedir. **Kural listesi** olarak da adlandırılabilen birleştirme matrisinin içindeki listelerde **kural kodları** bulunur. Bu kural kodları Şekil 3.7 ile görsel hale getirilen birleştirme matrisinde (kural matrisi) kutularla gösterilmiş; kodların irtibatlı olduğu birimlere ise oklarla işaret edilmiştir.



**Şekil 3.7: Birleştirme matrisi liste elemanları ve irtibatlı oldukları birimler**

Birleştirme matrisinde (Şekil 3.7) binlercesi bulunan kural kodları, çeviri modülünün HD dizgisi üretimi için ihtiyaç duyduğu dilbilimsel verileri tutan listelerdir. Her listede 4 ile 11 arası eleman bulunur. Bu liste elemanları, görevleri itibarıyla dörde ayrılırlar. Baştan iki ile beş arası kod, **kontrol kodlarıdır**. Onların sağında kural kodlarının merkezi kabul edilen **karar kodu** vardır. Karar kodunun sağında ise 2 ikame kodu (**Kod 1, Kod 2**) ile 2 ek dizgisi (**Ek 1, Ek 2**) elemanları sıralanır.

Tüm kural listesi elemanları (Şekil 3.7) birbirleriyle irtibatlı şekilde çalışır. Bu irtibat, karar kodu yardımı ile gerçekleşir. Elemanlar arasındaki bu ilişkinin doğru tarif edilebilmesi için tüm birleştirme matrisi ‘Kontrol kodları’, ‘Karar kodu’, ‘İkâme kodları’ ve ‘Ek dizgileri’ olmak üzere dört başlıkta ele alınmıştır.

### 3.2.2.1. Kontrol kodları

Ayrıştırma ünitesinde üretilen kod listesi ve Türkçe dizgi listesi ilk olarak çoklu kod sorgulama birimine gelir. Birim, listelerdeki hangi kod ve dizgiler üzerinde değişiklik (birleştirme, değiştirme, silme vb.) yapılacağını tespit eder. Bunun için aşağıdaki örnekte olduğu gibi önce kod listesi elemanlarının, (3.14) kural listelerinden birindeki (3.15) kontrol kodlarıyla eşleşmelerinin bulunması gerekir.

['\_ler', '', 'kapı', 'in', '\_dir', 'araba', 'eski', 'yaşlı', '-m', '|'], (3.14)

['PLR', 'TPS', 'N', 'S\_', 'BE', 'N', 'J', 'Ja', 'DT', '|']

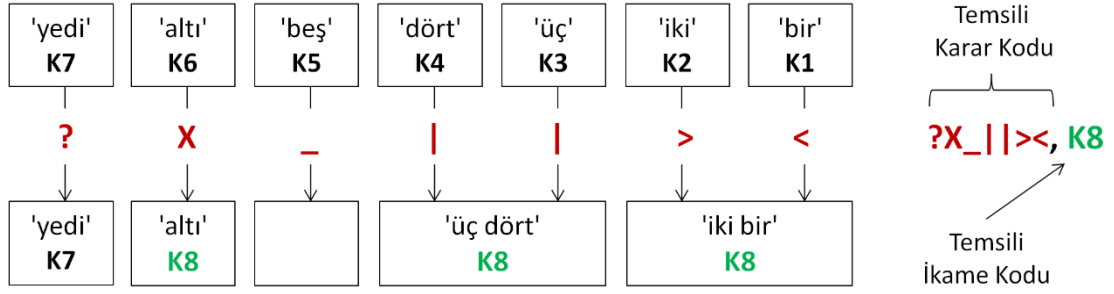
['N', 'S\_', 'BE', 'N', '?X\_?', 'S\_'] (3.15)

Örnekte (3.14) dört Türkçe dizgi listesi elemanı ['kapı', 'in', '\_dir', 'araba'] ile aynı indise sahip dört kod listesi elemanının, ['N', 'S\_', 'BE', 'N'] kural listesindeki (3.15) dört kontrol koduyla bire bir aynı olduğu görülür. Bu eşleşme, bulunan dört kod ve onlarla aynı indise sahip dizgiler (3.14) üzerinde değişiklik yapılacağı anlamına gelir. Ancak değişikliğin ne olacağı, kontrol kodunun sağındaki karar koduna (?X\_?) bağlıdır.

#### 3.2.2.2. Karar kodu

Karar kodu, solundaki kontrol kodlarının toplam sayısı kadar dizgi karakterine sahiptir. Her karakter, örnekte (3.15) olduğu gibi ('?'→'N', 'X'→'S\_', '\_'→'BE', '?'→'N') bir kontrol kodunu temsil edecek şekilde sıralanır. Hangi kod üzerinde ne tür bir ekleme-çıkarma yapılacağı karar kodu karakterlerinin sırasına ve ne olduklarına göre değişir. Her karar kodu karakteri bir kurala işâret eden komut karakteridir.

Şekil 3.8 ile tüm komut karakterlerinin toplu halde olduğu (temsili) bir karar kodu ve karar kodunun ihtiyaç duyduğu (temsili) bir ikame kodu sunulmuş; bu şekilde her komut karakterinin, kod ve dizgilerde yaptığı değişiklikler görsel hale getirilmiştir. Dizgilerde cümle yerine birden yediye kadar olan rakamlar kullanılmıştır ve bu rakam dizgilerini temsilen de K1, K2,.. K8 kodları tercih edilmiştir (Şekil 3.8).



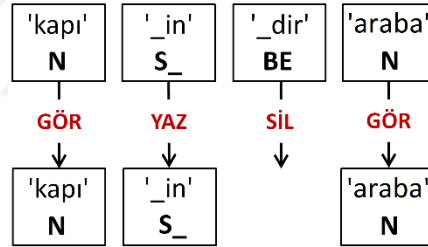
**Şekil 3.8: Komut karakteri yardımıyla kod ve dizgi üzerinde yapılan değişiklikler**

Tez modelinde önerilen çeviri modülünde, sunulan örnekteki (Şekil 3.8) gibi dizgi kodları, (K1, K2, K3,..) karar kodu (?X\_||><) veya ikame kodları (K8) bulunmaz. Tüm dizgi ve kodlar, daha iyi anlaşılabilmesi amacıyla bu örneğe mahsus olmak üzere gerçekteki hallerinden farklı (tüm yapısal değişikliklerin tek bir karar kodu üzerinden gerçekleştirildiği bir örnek ile) sunulmuştur.

Karar kodunun (Şekil 3.8) ilk komutu olan '?' ilk kutudaki 'K7' kodlu 'yedi' dizgisine bağlıdır. 'Gör' anlamında olan bu komut sadece kodun yerinde olup olmadığına bakar. Karar kodunun ikinci 'X' komutu, ikinci kutudaki 'K6' kodlu 'altı' dizgisine bağlıdır. 'Yaz' anlamında olan bu komut karakteri dizgiye dokunmamış; sadece ikame kodunu (K8) eski kodun yerine yazmıştır. Kodun üçüncü komutu olan '-' üçüncü kutudaki 'K5' kodlu 'beş' dizgisiyle irtibatlıdır. 'Sil' anlamına gelen bu komut hem dizgi hem de kodu siler. Dört ve üçüncü sıradaki '|' ve '|' komutları, dördüncü ve üçüncü kutulardaki 'dört' ve 'üç' dizgilerini 'üç dört' şeklinde tek dizgi haline getirir. Daha önceki 'K4' ile 'K3' kodlarının yerine ikame kodu (K8) getirilir. Bu komutlar 'ekle' anlamındadır ve her zaman bitişik (||) olurlar. İkinci ve birinci sırada olan '>' ve '<' komutları ise ikinci ve birinci kutulardaki 'iki' ve 'bir' dizgilerini 'bir iki' olarak değil; tersine 'iki bir' olarak tek dizgi yapar. Burada da 'K2' ile 'K1' kodlarının yerine ikame kodu (K8) getirilir. Bu karakterler (><) de her zaman bitişik olur ve 'ters ekle' (sağdakini sola, soldakini sağa geçir) anlamına gelir.

Verilen kural listesi örneğine (3.15) tekrar dönülecek olursa karar kodundaki (?X\_?) birinci ve dördüncü komut karakteri ‘?’ sadece birinci ve dördüncü kodun aranan listede olup olmadığını kontrol eder. (Kodlar üzerinde silme veya birleştirme gibi işlemler yapmaz.) Üzerinde değişiklik yapılması gereken kodlar, önlerindeki veya arkalarındaki kodlara göre tespit edilmelidir. Bunun için ‘?’ (gör) karakteri kullanılır. Karar kodundaki (3.15) ikinci komut karakteri ‘X’, ikinci kodun yerine (bu örnekte aynı) ikame kod geleceğini haber verir. Kodun üçüncü sırasında görülen komut karakteri ‘\_’ ise üçüncü sıradaki kodun ve bağlı olduğu dizginin kaldırılacağı anlamındadır.

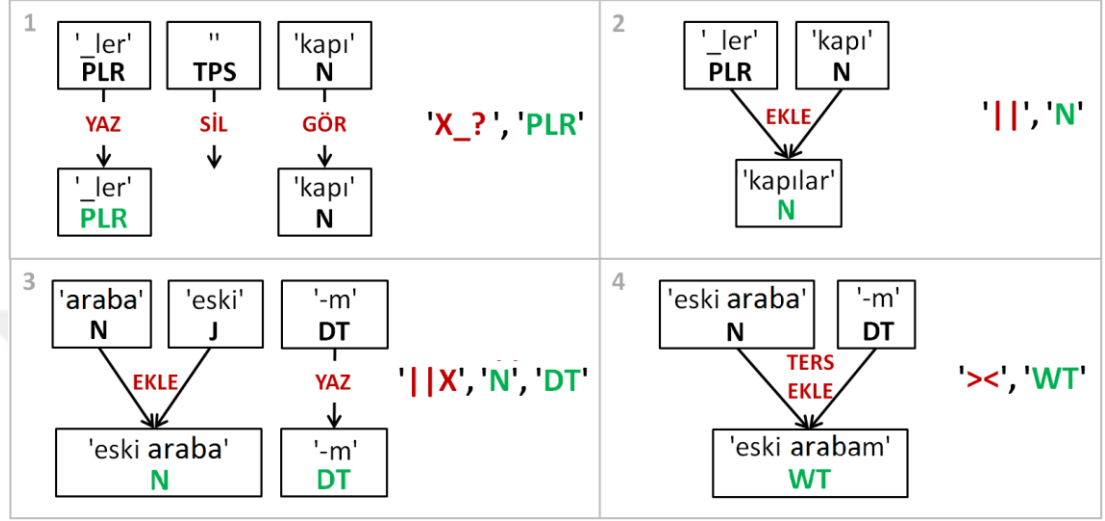
Netice itibarıyla örnekteki (3.15) ‘?X\_?’ liste elemanı, “gör, yaz, sil, gör” komutları sıralamasına sahip bir karar kodudur. Bu komut sıralamasına göre, kodlar ve dizgilerde yapılacak olan değişiklikler ise Şekil 3.9 ile aşağıda görselleştirilmiştir.



**Şekil 3.9: Karar kodu (?X\_?) uygulama örneği**

Ayrıştırma ünitesi çıktısında (3.14) geçen “...car’s door...” (...arabanın kapısı...) ifadesinin dizgilere bölünmüş hali ‘car’, ‘s’, ‘door’ şeklindedir. Yukarıdaki örnekte (Şekil 3.9) bu dizgilere denk gelen Türkçe dizgi ve kodlar sunulmuştur. Kutulardan iki ve üçüncü sıradakiler ‘s’ dizgisinin karşılığı olan iki adet dizgi kodu (‘S\_’, ‘BE’) ve iki adet Türkçe dizgidir (‘-in’, ‘-dir’). Burada modül, baştaki ve sondaki kodlara bakarak ‘s’ dizgisinin karşılığı olan iki koddan (‘S\_’, ‘BE’) uygun olanını tercih etmiş ve ‘BE’ kodunu irtibatlı olduğu ‘-dir’ dizgisiyle birlikte silmiştir (Şekil 3.9). Çünkü dizgi (-’s) olmak anlamındaki “is” kelimesin kısaltılmışı (-’s) olarak değil; sahiplik ifade eden yapıda kullanılmıştır.

İçerdiği komut karakterlerine göre fonksiyonları belirlenen karar kodlarının farklı uygulama örnekleri için Şekil 3.10 incelenebilir.



**Şekil 3.10: Karar kodları uygulama örnekleri**

Yukarıda (Şekil 3.10) karar kodu işletilmiş olan dört farklı uygulama örneği sunulmuştur. (Karar kodlarının sağında ikame kodları görülmektedir.) Kutulardaki liste elemanları (kod ve dizgiler) sistemde geçerli olan kod analizi sırasına göre (sağdan sola) dizilmiştir.

Şekil 3.10 ile gösterilen birinci kutuda 'X\_?' (yaz, sil, gör) karar kodu kullanılmış ve bir önceki görselde (Şekil 3.9) olduğu gibi, silme yapılmıştır. İkinci kutunun karar kodu '|' (ekle) seçildiği için kod ve dizgiler birleşerek 'N' ikame kodlu 'kapılar' dizgisini oluşturmuştur. Üçüncü kutudaki '|X' (ekle, yaz) kodu, 'eski' ile 'araba' dizgilerini 'N' ikame koduyla birbirine eklemiş (eski araba) ve bu kodların önündeki 'DT' kodunu da değiştirmeyip aynı ikame kodunu yazmıştır. Dördüncü karar kodundaki '<' (ters ekle) karakterleri ise ikame kodu olan 'WT' yi dizgiye vermeden önce, baştaki 'DT' kodlu '-m' ekini 'N' kodlu 'eski araba' dizgisinin sonuna eklemiştir (Şekil 3.10).

### 3.2.2.3. İkâme kodları

Karar kodunda bulunan 'X' (yaz), '||' (ekle) ve '><' (ters ekle) karakterleri, değiştirmek durumunda oldukları kodların yenilerini ikame kodlarından temin eder. Bu iki (bazen üç) ikame kodunun kural kodlarındaki yeri, kural listesindeki karar kodu ile ek dizgileri (Ek 1, Ek 2) arasındadır (Şekil 3.7).

Görseldeki (Şekil 3.10) üçüncü kutuda bulunan karar kodu, hem ekleme hem de yazma işinde görevlidir ve bu kod (||X) birinci değişiklik olan ekleme (||) için birinci ikame (Kod 1) kodunu, ikinci değişiklik olan yazma (X) için de ikinci ikame (Kod 2) kodunu kullanmıştır.

### 3.2.2.4. Ek dizgileri

Bu iki dizgi, (Ek 1, Ek 2) Türkçede mevcut olduğu halde İngilizcede yapısal karşılığı olmayan ekler için kullanılır. İngilizcede üç dizgiden ('house', 's', 'cat') oluşan "house's cat" ifadesi Türkçeye dört dizgiyle ('ev', '-in', 'kedi', '-si') "evin kedisi" şeklinde çevrilir. Yani isim tamlamalarındaki tamlanan (iyelik) eki İngilizce isim tamlamalarında bulunmaz. İngilizce dilinde iki ismin yan yana gelmesi bir tamlama oluşturmak için yeterlidir. Bu nedenle "evin kedisi" ifadesindeki ikinci kelime (kedi) için ihtiyaç duyulan '-si' dizgisi, ek dizgi elemanlarından ikincisi olan 'Ek 2'den temin edilir. Birinci kelimedeki olabilecek ek ilaveleri için de 'Ek 1' kullanılır: "See Istanbul." cümlesi, birinci kelimeye '-u' eki ilave edilerek "İstanbul+u gör." şeklinde çevrilir.

### 3.2.3. Kod ve dizgi üretim

Karar koduna göre üzerinde değişiklik yapılmasına karar verilen iki kod ve onlara (indis no ile) bağlı olan iki Türkçe dizginin birleşme, değişme veya silinme işlemleri 'kod ve dizgi üretim' biriminde tanımlanmış olan birleştirme metodu yardımıyla gerçekleşir. Söz gelimi üç kod içerisinde ilk ikisinin birleşmesini ve sonraki kodun da değişmesini isteyen bir karar kodu (||X) görüldüğünde bu metod, birleşecek olan kodların yerine 'Kod 1'i; değişecek olan kodun yerine de 'Kod 2'yi geçirir. Tam tersi, üç kod içerisinde ilkinin değişmesini ve sonraki iki kodun birleşmesini isteyen

bir karar kodu (X||) görüldüğünde de değişecek olan kodun yerine ‘Kod 1; birleşecek olan kodların yerine de ‘Kod 2 getirilir.) Kodlar üzerinde yapılan her değişiklik, kodların (indislerle) bağlı olduğu dizgilere de yapılır. Ancak değişiklik yapılan dizgiler dizgi listelerine geri dönmeden önce yapım-çekim eki ilave bölümünde, birleştirme matrisinden gelen (varsa) ekleri (Ek 1 ve/veya Ek 2) almak durumundadır.

#### **3.2.4. Yapım-çekim eki ilave**

Neredeyse tüm biçimbirimlerin, kullanılan sesli harf ve biçimbirim sınırlarındaki sessiz harf yönünden farklı şekilleri bulunmaktadır. Örneğin; “paket+ten” kelimesinde ekin ilk sessizi ve seslisi birleştiği gövdenin son sessiz ve seslisi ile uyum içinde olmak için ‘t’ ve ‘e’ olarak seçilmiş; “araba+dan” kelimesi ise bir sesli harf ile bittiği için ekin ilk sessiz harfi ‘d’ olarak kalmıştır. Ancak sesli harf uyum için *a* olmak durumundadır. Sesli (ünlü) uyumu soldan sağa silsile şeklinde gelişmektedir (Oflazer, 2016). Biçimbirimler düzeyinde gelişen bu durum için birleştirme ünitesi sınıfı içerisinde bir ek-ilave metodu tanımlanmıştır. Ancak bu metod çoğu kelimelerin eklerle birleştirilmesinde başarılı olsa da bazı ek dizgilerine mahsus istisnalar olabilmektedir. “bil+ir” ve “sil+er” kelimelerinde ‘-ir’ ve ‘-er’ ekleri bu duruma örnek gösterilebilir. (Bu tür istisnai durumların önüne geçebilmek için modül çıktılarında gerekli düzeltmeleri gerçekleştiren bir metod çalışmaya eklenmiştir.)

Birimde tanımlanmış olan ek-ilave metodu yalnız ses uyumlarıyla değil; ismin yalın, belirtme, yönelme, bulunma ve ayrılma hâllerine göre değişen eklerle de ilgilenmek durumundadır. Tablo 3.1 ile verilen örneklerde aynı KD isim dizgisinin (...*home*) farklı fiillerle kullanıldığında, HD çevirisinde farklı ekler aldığı görülmektedir.

**Tablo 3.1: Türkçede fiillere göre değişen ismin beş hâli**

Örn.	KD Cümlesi	HD Cümlesi	HD İsim Hâlleri
1	<i>Say home.</i>	Ev de.	Yalın hâli
2	<i>Clean home.</i>	Ev+i temizle.	'-i' hâli (Belirtme)
3	<i>Go home.</i>	Ev+e git.	'-e' hâli (Yönelme)
4	<i>Stay home.</i>	Ev+de kal.	'-de' hâli (Bulunma)
3	<i>Leave home.</i>	Ev+den ayrıl.	'-den' hâli (Ayrılma)

Tablodaki (Tablo 3.1) tüm kaynak dil cümlelerinde '*home*' kelimesi herhangi bir kelime veya ek ilave edilmeden sadece farklı fiillerle kullanılmıştır. Fakat bu cümlelerin çevirisi olan Türkçe cümlelerde 'Ev' kelimesi yalın durum hariç dört ayrı hâl eki ('-i', '-e', '-de', '-den', ) olarak kullanılmıştır.

Ek-ilave metodu bu tip farklı fiiller kullanılan çevirilerde, isme uygun hal ekleri ilave edebilecek şekilde tasarlanmıştır. Sözlük aşamasında her fiil veya isim farklı kodlar almış; birleştirme analizlerinde ise bu kodlar üzerinden çözümlenmeler yapılarak gerekli eklerin ilave edilmesi sağlanmıştır.

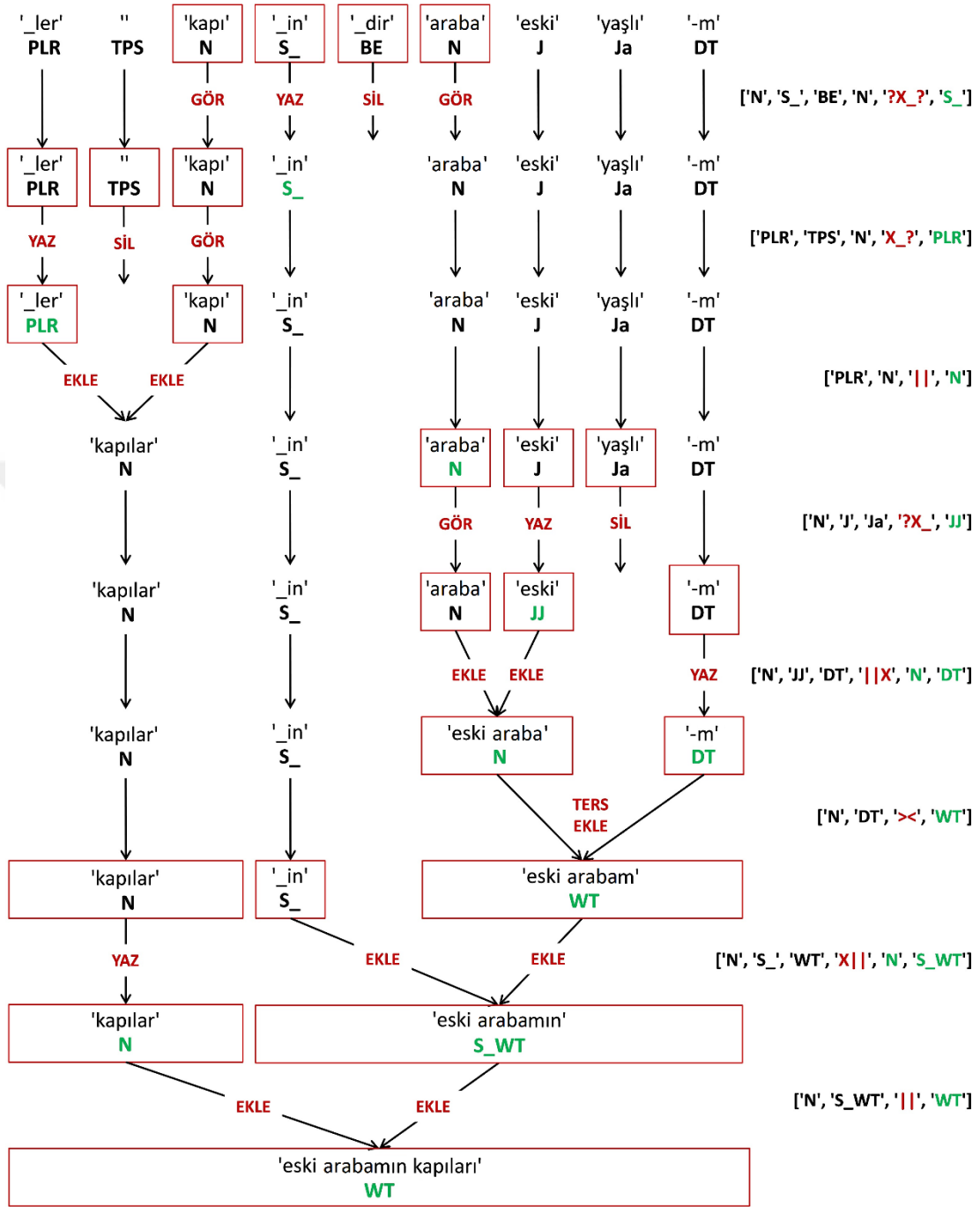
### 3.2.5. Birleştirme ünitesinde HD üretim aşamaları

Buraya kadar bahsi geçen tüm birleştirme ünitesi birimlerinde gerçekleşen üretim aşamalarının, kod ve dizgi listeleri ile görsel hale getirilmesi yararlı olacaktır. (Şekil 3.11). Görselde her üretim aşaması için gerekli olan kural kodlarına ve bu kodlar yardımıyla yapılan değişikliklere (silme, değiştirme, ekleme vb.) yer verilmiş; devam eden paragraflarda ise her bir üretim aşaması tek tek ele alınarak tarif edilmeye çalışılmıştır.

Şekildeki (Şekil 3.11) dizgilerin hemen altında bulunan dizgi kodları (*tokens*) 'DT', 'Ja', 'J', 'JJ', 'N', 'BE', '\_S', 'TPS', 'PLR', 'WT', 'S\_WT' şu şekilde tarif edilebilir: (i) 'My', 'Your', 'The' gibi niteleyici kelimeler (*Determiners*) için 'DT'; (ii) hangi anlamının kullanılacağı seçilmemiş sıfatlar (*Adjectives*) için 'J' veya 'Ja'; (iii) hangi anlamının kullanılacağı belirlenmiş sıfatlar (*Adjective*) için 'JJ'; (iv) isimler (*Nouns*)

için 'N'; (v) 'olmak' fiili için 'BE'; (vi) sahiplik bildiren 's' eki (*Possessive*) için 'S\_'; (vii) geniş zamanda üçüncü tekil şahıslarda (*Third Person Singular*) kullanılan 's' takısı için 'TPS'; (viii) isimleri çoğul (*Plural*) yapan 's' takısı için 'PLR'; (ix) isim ifadeleri (*Noun phrases*) için WT; (x) sahiplik (*Possessive*) eki 's' bulunduran isim ifadeleri (*Noun phrases*) için ise 'S\_WT' kullanılmıştır.





Şekil 3.11: Kontrol kodları ve HD dizgisi üretimi aşaması

Görselin (Şekil 3.11) en üstünde sıralanan kod ve dizgiler “my old car’s doors” dizgisinin ayrıştırma ünitesi çıktısıdır (3.14). Kutulardaki dizgilere bağlı dizgi kodlarının, görselin sağında her üretim aşamasında kullanılmış olan kural kodlarındaki kontrol kodları ile eşleştiği görülmektedir. Aşağı yönlü oklar üretim

aşamalarını gösterir. Ok üzerindeki yönlendirmeler (gör, yaz, sil vb.) kural kodlarındaki karar koduna (kontrol kodlarının sağında) aittir. Karar kodunun sağındaki ikame kodlarının ise değişiklikleri yapılan kutulardaki kodların yerine kullanıldıkları görülmektedir. (Kural kodları için ‘Birleştirme matrisi’ başlığı (3.2.2) incelenebilir.)

Görseldeki (Şekil 3.11) dizgi kodları üzerinde uygulanan kural listeleri, uygulanma önceliklerine göre aşağıda (3.16) - (3.23) sıralanmıştır.

['N', 'S\_', 'BE', 'N', '?X\_?', 'S\_', "", "", ""], (3.16)

['PLR', 'TPS', 'N', 'X\_?', 'PLR', "", "", ""], (3.17)

['PLR', 'N', "", '||', 'N', "", "", ""], (3.18)

['N', 'J', 'Ja', '?X\_', 'JJ', "", "", ""], (3.19)

['N', 'JJ', 'DT', '||X', 'N', 'DT', "", ""], (3.20)

['N', 'DT', '><', 'WT', "", "", ""], (3.21)

['N', 'S\_', 'WT', 'X||', 'N', 'S\_WT', "", ""], (3.22)

['N', 'S\_WT', "", '||', 'WT', "", "", '--si'], (3.23)

Yukarıda sunulan listeler (3.16) - (3.23) birleştirme matrisinde bulunan binlerce liste içerisinde sadece “*my old car’s doors*” dizgisine ait kod listesi (3.24) üzerinde uygulanmış kural listeleridir. Kod ve dizgi listelerinden (3.24) HD dizgisinin üretilmesi aşamaları bu kural listeleri (3.16) - (3.23) üzerinden anlatılacaktır.

Aşağıdaki (3.24) iç içe liste, ayrıştırma ünitesinden gelen kod ve dizgilerdir. (KD dizgisinin, HD kelime ve kodlara ayrılmış olan çıktısıdır.) Kural kodları içindeki (3.16) kontrol kodlarıyla eşleşmiş olan kodlar (3.24) için kalın harf karakterleri kullanılmıştır. Eşleşen kodlar üzerinde yapılacak her değişiklik, aynı indis değerine sahip (yine kalın olarak yazılmış) dizgilere (3.24) de uygulanır.

[['PLR', 'TPS', 'N', 'S\_', 'BE', 'N', 'J', 'Ja', 'DT', '|'], (3.24)

['\_ler', '', 'kapı', 'in', '\_dir', 'araba', 'eski', 'yaşlı', '-im', '|']]

Dizgi kodları (3.24) üzerinde birinci kural (3.16) işletildiğinde aşağıdaki kod ve dizgi listeleri (3.25) elde edilir. Kod ve dizgilerin yeniden üretildiği bu çözümlenme aşamasında karar verici konumda olan karar kodu, (?X\_?) üzerinde işlem yapılmış kod listesinin sağına (3.25) eklenmiştir. Karar kodu komut karakterlerine göre kod ve dizgi üzerinde silme işlemi uygulanmış; 'BE' kodu ve ona bağlı '-dir' eki buldukları listelerden silinmiştir.

[['PLR', 'TPS', 'N', 'S\_', 'N', 'J', 'Ja', 'DT', '|'] ?X\_? (3.25)

['\_ler', '', 'kapı', '\_in', 'araba', 'eski', 'yaşlı', '-im', '|']]

Kodlarda (3.25) ikinci kural (3.17) uygulandığında aşağıda görülen listelerin (3.26) üretimi gerçekleşmiş olur. Karar kodu (X\_?) komut karakterlerine göre kod ve dizgilerde yine silme işlemi uygulanmış; 'PLR', 'TPS', 'N' kodları içinden 'TPS' ve ona bağlı boş liste karakteri (") silinmiştir. (TPS kodu dizgisinin boş olması sadece ek ilavelerinde kullanıldığı içindir. Silinmeyip de kullanılması durumunda boş olan liste elemanının yerine kural listesinden [Ek1 veya Ek2'den] bir ek gelecektir.)

[['PLR', 'N', 'S\_', 'N', 'J', 'Ja', 'DT', '|'] X\_? (3.26)

['\_ler', 'kapı', '\_in', 'araba', 'eski', 'yaşlı', '-im', '|']]

Yukarıdaki kodlarda (3.26) sıradaki kural (3.18) uygulanmış ve aşağıda görülen listelerin üretimi (3.26) gerçekleşmiştir. Karar kodu (||) komut karakterlerine göre kod ve dizgilerde birleştirme yapılmıştır. 'PLR' ve 'N' kodları yerine 'N' ikame kodu yazılmış ve irtibatlı oldukları liste karakteri '\_ler', 'kapı' dizgileri de bir araya getirilerek 'kapılar' dizgisi elde edilmiştir. ('kapı' ve '-lar' dizgileri, yapım-çekim eki ilave birimi metodu yardımıyla 'kapı+lar' şeklinde birleşmiştir.)

['N', 'S\_', 'N', 'J', 'Ja', 'DT', '|'] || (3.27)

['kapılar', '\_\_in', 'araba', 'eski', 'yaşlı', '-im', '|']

Üretilen liste kodlarında (3.27) bir sonraki kural (3.19) uygulanmış ve aşağıdaki listeler (3.28) elde edilmiştir. Karar koduna (?X\_) göre kod ve dizgilerde silme yapılmış; 'J' ve 'Ja' kodları arasında tercih yapılarak yerlerine 'sıfat tercihi yapıldı' anlamına gelen 'JJ' ikame kodu getirilmiştir. Silinen 'J' kodu ile aynı indisteki 'yaşlı' dizgisi de listelerden kaldırılmıştır.

['N', 'S\_', 'N', 'JJ', 'DT', '|'] ?X\_ (3.28)

['kapılar', '\_\_in', 'araba', 'eski', '-im', '|']

Üzerinde silme işlemi gerçekleştirilen kod listesinde (3.28) bir sonraki kural (3.20) uygulanmış ve sıradaki listelerin (3.29) üretilmesi sağlanmıştır. Komut karakterlerine (||X) göre 'N' ve 'JJ' kodlarının, bağlı oldukları dizgilerle 'N' ikame kodunda birleşmesine; 'DT' kodu ve ona bağlı dizginin ise olduğu gibi kalarak üretilen listeye geçirilmesine karar verilmiştir.

(Kural listelerinde '||?' (ekle, gör) şeklinde bir karar kodu bulunmaz. Yerine '||X' (ekle, yaz) kullanılır. Böylece ekleme sonrası değiştirme gerektiren durumlar için iki ayrı karar kodu kullanılmayarak analiz sürecinde daha kestirme bir yol takip edilmiş olur.)

['N', 'S\_', 'N', 'DT', '|'] ||X (3.29)

['kapılar', '\_\_in', 'eski araba', '-im', '|']

Ortaya çıkan yeni liste kodlarına (3.29) sıradaki kural (3.21) uygulanmış ve aşağıdaki listeler (3.30) elde edilmiştir. Karar koduna (><) göre kod ve dizgilerde baştaki öne, öndeki başa gelecek şekilde birleştirme yapılmıştır. Yani baştaki 'DT' ve dizgisi 'N'nin sonuna gelecek şekilde birleşir. (Kodlar sağdan sola doğru dizildiği için 'DT' kodu 'N' kodunun önündedir.)

['N', 'S\_', 'WT', '|'] >< (3.30)

['kapılar', '\_\_\_in', 'eski arabam', '|']

Birleştirme işlemi gerçekleştirilen kod listesinde (3.30) müteakip kural (3.22) uygulanmış ve aşağıdaki listeler (3.31) üretilmiştir. Komut karakterlerine (X||) göre 'S\_' ve 'WT' kodları, bağlı oldukları dizgilerle 'S\_WT' ikame kodunda birleşmesine; 'N' kodu ve ona bağlı dizginin de olduğu gibi kalmasına karar verilmiştir.

['N', 'S\_WT', '|'] X|| (3.31)

['kapılar', 'eski arabamın', '|']

Birleştirme işlemi yapılmış kod listesinde (3.31) son kalan kural (3.23) da işletildikten sonra HD dizgi kodu ve HD Türkçe dizgisini içeren listeler (3.32) üretilmiş olur. Komut karakterlerine (||) göre 'N' ve 'S\_WT' kodları ve bağlı oldukları dizgiler birleşmiş; dizgi kodu olarak 'WT' ikame kodu seçilmiştir. Ancak hemen öncesinde kural kodundaki (3.23) 'ek 2'de bulunan '-si' eki, ikinci sıradaki 'kapılar' dizgisine (3.32) ilave edilmiş; yapım-çekim eki metodu sayesinde birleşeceği dizgiye (ilave edilmeden önce harf uyumuna göre değişerek) 'kapılar+' şeklinde eklenmiştir.

['WT', '|'] || (3.32)

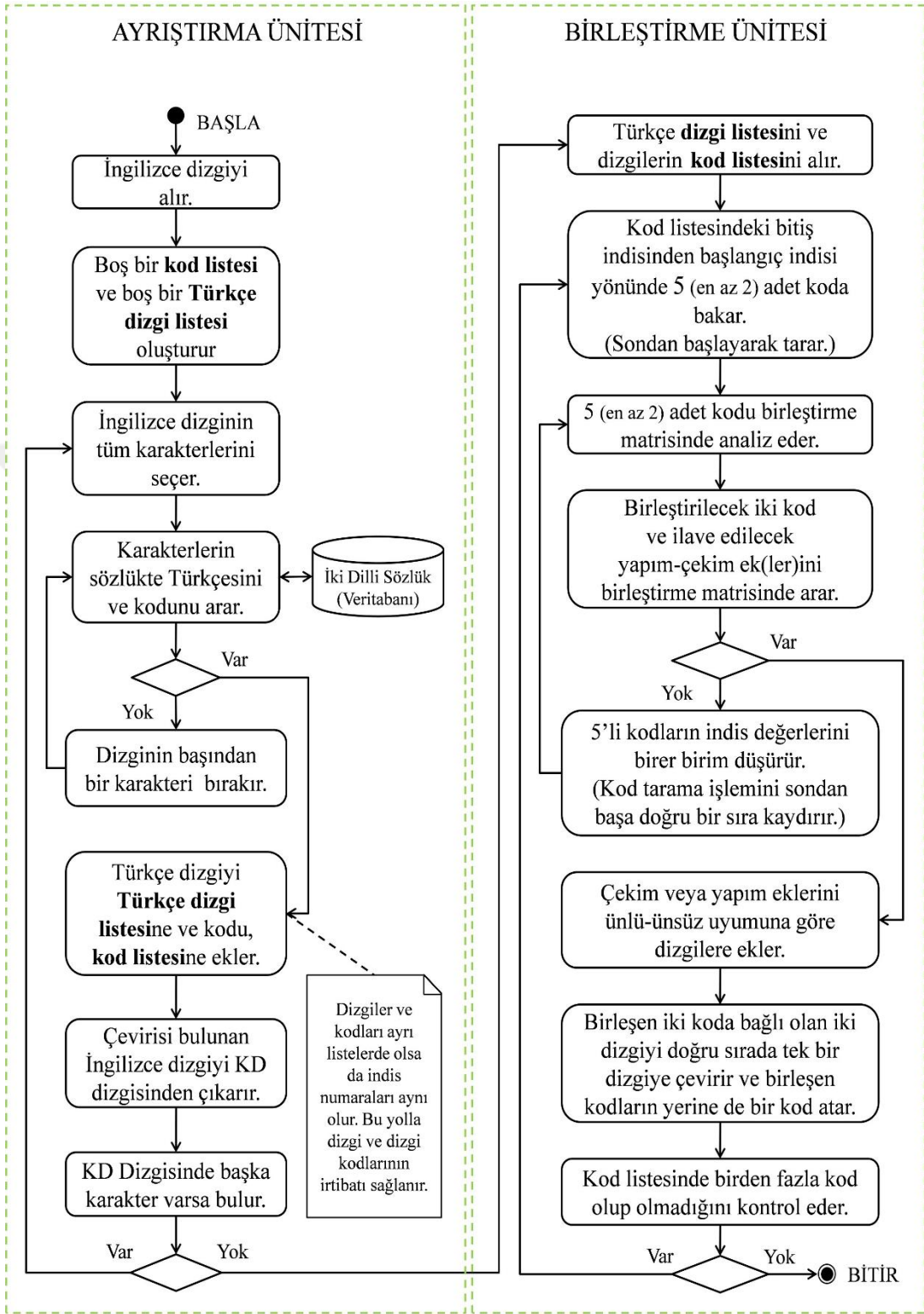
['eski arabamın kapıları', '|']

Elde edilen HD dizgi kodu 'WT' ile 'eski arabamın kapıları' dizgisinin (3.32) sağında bulunan '|' karakterinin, en az iki elemana ihtiyaç duyan birleştirme matrisi analizleri için listelerde bulunması gereklidir. Dizgi listesinde cümle başı anlamına gelen bu dizgi karakteri HD dizgi çıktısına (3.33) dâhil edilmez.

['eski arabamın kapıları'] (3.33)

Özetlemek gerekirse bu aşamaya kadar KMMÇ modülünün iki ana birimi olan ayrıştırma ve birleştirme üniteleri ele alınmış; (i) ayrıştırma ünitesi başlığı altında KD girdisinin iki dilli sözlük yardımıyla gerekli ek, kelime veya ifade dizgilerine (*strings*) ayrıldığı; bu dizgilerin kodlarla (*nonterminals* [devamlılar]) eşleştirildiği; eşleşen kod ile dizgi listelerinin birleştirme ünitesine gönderildiği ve (ii) birleştirme ünitesi başlığında bu kodların çözümlenmesi yoluyla dizgilerde hangi değişikliklerin yapılarak HD çıktısının üretildiği tarif edilmeye çalışılmıştır.

Yukarıda bahsi geçen tüm bu sürecin tek bir görsel üzerinde bir arada görülmesinde fayda olacağı düşüncesiyle KMMÇ modelinin iki ana birimi olan ayrıştırma ve birleştirme ünitesinde gerçekleşmekte olan üretim aşamaları hazırlanan aktivite diyagramı (Şekil 3.12) ile aşağıda sunulmuştur. Ayrıştırma ve birleştirme ünitelerinin özeti mahiyetinde olan bu diyagramda (Şekil 3.12) dizgi birleştirme işleminin (silme, değiştirme vb. hariç) gerçekleştirildiği üretim aşamaları görsel hale getirilmiştir. Müteakip iki paragrafta ise KD girdisi ve HD çıktısı arasında gerçekleşen üretim süreci bu görsel (Şekil 3.12) üzerinden tarif edilmiştir.



Şekil 3.12: Ayrıştırma ünitesi ve birleştirme ünitesi aktivite diagramı

Ayrıştırma ünitesine (Şekil 3.12) İngilizce girdi dizgisi alınır. Ayrıştırılacak kod ve dizgiler için boş kod listesi ve boş Türkçe dizgi listesi tutan bir iç içe liste oluşturulur. Üniteye alınan İngilizce dizgi karakterlerinin tümü iki dilli sözlükte aranır. Karakterler sözlükte bulunamadığında dizginin başından bir karakter eksiltilerek yeniden sözlük araması yapılır. Dizgi başından her seferinde bir karakter eksiltilerek yinelenen arama işlemi dizginin sözlükte eşleşeni bulunana kadar devam eder. Bulunan dizginin karşılığı olan Türkçe dizgi ve dizgi kodu sözlükten temin edilir. Türkçe dizgi, (daha önce oluşturulan) Türkçe dizgi listesine; dizgi kodu da (daha önce oluşturulan) dizgi listesine gönderilir. Listeye gönderilen dizgi, üzerinde arama yapılan KD dizgisinden çıkarılır. Kalan KD dizgisi için, başından bir karakter eksiltirerek yapılan yinelemeli sözlük taramasına devam edilir. KD dizgisinde tek bir karakter kalmayana kadar sözlükte kod ve Türkçe dizgi arama işlemi devam eder. Her döngüde sözlükten alınan kod ve Türkçe dizgi elemanları aynı indis değerine sahip şekilde listelerine eklenir. Bu indisler sayesinde kod ile dizgisi arasındaki irtibat sağlanmış olur. Oluşan bu listeler bir sonraki aşama olan birleştirme ünitesine gönderilir.

Birleştirme ünitesi (Şekil 3.12) gelen kod ve Türkçe dizgi listelerini (iç içe liste yapısında) aldıktan sonra tüm analizlerini (en sondan başlayarak) kodlar üzerinden yapar. Kodlarda yapılacak her değişiklik dizgi elemanları için de geçerli olur. Değişiklik için, kod listesindeki bitiş indisinden başlangıç indisine doğru 2 ile 5 arasında değişen sayıda kod seçilerek birleştirme matrisinde analiz edilir. Analiz, seçilen kodlar üzerinden birleşmesi gereken dizgiler ve eklerin aranması işlemidir. Birleşmesi gereken kodlar bulunamadığında seçilen kodlar sondan başa doğru bir sıra kaydırılır (Soldan bir kod çıkarılıp sağdan bir kod eklenir.) ve tekrar birleştirme matrisinde arama yapılır. Tüm kodlar bulunana kadar bu döngü devam eder. Birleşecek kodlar (ve bu kodlara bağlı olan Türkçe dizgiler) bulunduktan sonra eklenmesi gereken ekler (varsa) ilave edilir. Ekleri (varsa) ilave edilmiş iki dizgi doğru sıra ile arada bir boşluk karakteri olacak şekilde birleştirilir. Birleşen dizgilere ait eski iki kodun yerine de yeni bir (ikame) kod getirilir. Böylece hem kod hem dizgi listelerinde birer eleman eksilmiş olur. Birleştirme işlemi, kod listesinde birleştirilecek kod kalmayana kadar devam eder.

### 3.2.6. Üretim aşamalarında eş yazımlı kelimeler

Dilbilimde eş adlı, eş sesli ve eş yazımlı olarak bilinen kelimeler bazen birbiriyle karıştırılabilmektedir (Samedova, 2008). Bu nedenle mezkur kavramlar hakkında kısa da olsa bazı açıklamalar yapmak yerinde olacaktır. Ayrıca; açıklamalarda, tez içeriğine uygun olarak, İngilizce kelimeler üzerinden örneklendirmelerin yapılmasının daha açıklayıcı olacağı düşünülmüştür.

#### 3.2.6.1. Eş adlı, eş sesli ve eş yazımlı kavramı:

Eş adlılar (*Homonyms*) farklı anlamlara geldiği halde aynı şekilde telaffuz edilen kelimelerdir (A.g.e.). Örneğin; *'fast'* kelimesi /fæst/ (feest) olarak okunur ve hem “oruç tutmak” anlamında bir fiil hem de “hızlı” anlamında bir sıfattır. Eş sesliler (*Homophones*) farklı anlamlara geldiği gibi farklı yazılıp aynı şekilde telaffuz edilen kelimelerdir. Bunlara, “deniz” ve “görmek” manalarına gelen ve /si:/ (sii) olarak okunan *'sea'* ve *'see'* kelimeleri örnek gösterilebilir. Eş yazımlılar (*Homographs*) ise birden fazla anlama sahip olduğu halde yazımlarının aynı olduğu (Vardar, 2002) kelimelerdir ve bu kelimelerin eş yazımlı kabul edilebilmesi için (*Heteronym* olarak bilinen kelimelerdeki gibi) farklı telaffuzlarının olması şeklinde bir şart yoktur (Sakar, 2021). Buna örnek olarak *'object'* ve *'type'* kelimeleri gösterilebilir. Birinci kelimenin “nesne” anlamında kullanılan /'ɑ:bdʒekt/ ('abcekt) ve “itiraz etmek” anlamında kullanılan ise /əb'dʒekt/ (ıb'cekt) şeklinde telaffuz edilir. İkinci kelimenin anlamları ise “tip” ve “yazmak”tır ve telaffuzu /taɪp/ (tayp) şeklindedir.

Tez kapsamında kullanılan çeviri değerlendirme ölçeği (Ek: 1 – Ek: 10) için tercih edilen cümleler incelendiğinde, yukarıda bahsedilen kavramlar arasında eş yazımlılık kavramının tercih edilmesinin daha kapsayıcı olacağı düşünülmüştür. Çünkü yazımı aynı olup söylenişi farklı (*Heteronym*) kelimeler ile yazımı aynı ve söylenişi de aynı (*Homonym*) kelimeler için, eş yazımlı (*homograph*) tabiri bir üst kavram niteliğindedir.

### 3.2.6.2. Eş yazımlı bulunduran dizgi çevirileri:

Eş yazımlı kelimelerin bulunduğu cümle çevirilerinin günümüz makine çevirisi teknolojisinde dahi doğru olarak yapılabildiği söylenemez. Özellikle, uzun mesafe bağımlılıkları veya İngilizce-Türkçe gibi uzak yapılara sahip diller söz konusu olduğunda bu tür metin çevirilerinin çok daha zor olduğunu eklemek gerekir.

Bu tez çalışmasında önerilen KMMC modülünde eş yazımlı kelime çevirilerindeki sorun, dilbilimsel verinin tutulduğu birleştirme matrisinde (3.2.2) yapılan analizler yoluyla halledilmeye çalışılmaktadır. Bu analizlerden birine örnek niteliğinde aşağıda (3.34) iki eş yazımlı kelime içeren “*Old fly.*” dizgisinin ayrıştırma ünitesi çıktısı (birleştirme ünitesi girdisi) sunulmuştur.

[['V', 'V1', 'NNa', 'J', 'Ja', '|'],  
['uç', 'uçur', 'sinek', 'eski', 'yaşlı', '|']]

(3.34)

Ayrıştırma ünitesine iki kelimedenden oluşan bir dizgi girildiği halde ünitenin çıktısında (3.34) 5’er elemanlı (‘|’ hariç) kod ve dizgi listelerinin olması, ‘*old*’ (eski, yaşlı) kelimesinin iki; ‘*fly*’ (uçmak, uçurmak, sinek) kelimesinin ise üç farklı anlamda sözlük girdisi bulunmasından kaynaklanır. Ayrıştırma ünitesi, tek bir eş yazımlı kelimenin beş farklı anlamını sözlükten bularak, dizgi kodları ile birlikte, birleştirme ünitesine gönderebilir. Bundan sonrası, birleştirme ünitesinde uygulanacak olan analizlerin işidir.

Verilen örneğin (3.34) analizinde kullanılmış olan kural listeleri uygulanma sıralarına göre aşağıda (3.35), (3.36), (3.37), (3.38) verilmiştir.

['V', 'V1', 'NNa', 'J', 'X\_??', 'V', "", "", "", ""],

(3.35)

['V', 'NNa', 'J', '\_X?', 'Na', "", "", "", ""],

(3.36)

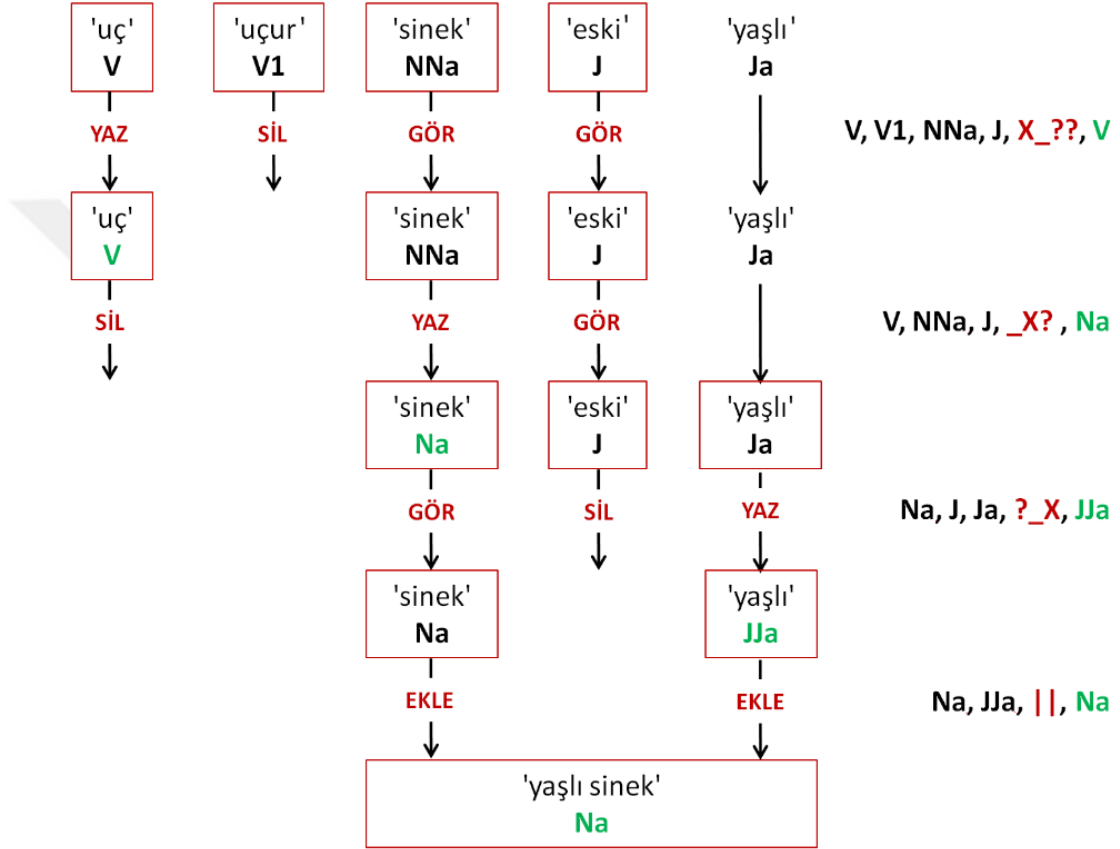
['Na', 'J', 'Ja', '?\_X', 'JJa', "", "", "", ""],

(3.37)

['Na', 'JJa', "", '|', 'Na', "", "", "", ""],

(3.38)

Bu listeler (3.35) - (3.38) birleştirme matrisinde bulunan binlerce liste içinden yalnız “Old fly.” dizgisine ait kod listesi (3.34) üzerinde uygulanmış kural listeleridir. Dizgiye ait analizler aşağıda (Şekil 3.13) görüldüğü gibi şeklin sağındaki kural listelerindeki kırmızı olarak görülen kural kodlarına göre yapılır.

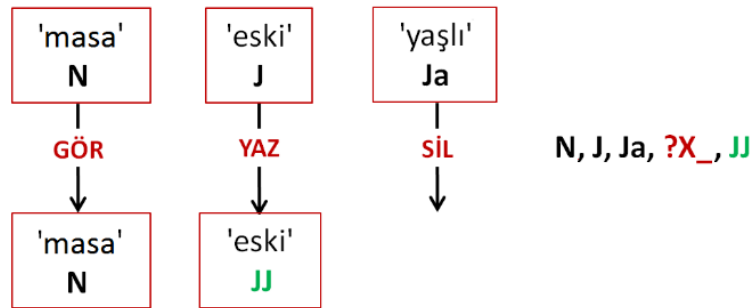


**Şekil 3.13: KMMC’de eş yazımlı bir kelime için doğru anlam seçimi analiz aşamaları**

Yukarıdaki dizgi analizleri (Şekil 3.13) kural listeleri (3.35) - (3.38) yardımıyla dört aşamada gerçekleşmiştir: (i) Birinci kuralın (3.35) işletildiği ilk aşamada kırmızı kutular ile gösterilen dört kod analiz edilmiş; 'J' (sıfat kodu) ile tanımlanan kelimedenden sonra (solunda) doğal dilin kuralı gereği bir isim yapısının bulunması beklendiği için 'NNa' (isim kodu) bırakılarak 'V1' (fiil kodu) ve ona bağlı 'uçur' dizgisi silinmiştir. (ii) İkinci kuralın (3.36) işletildiği ikinci aşamada kırmızı kutular ile gösterilen üç

kod analiz edilmiş ve yine 'J' (sıfat kodu) ile tanımlanan kelimedenden sonra fiil gelemeyeceği varsayımı ile 'NNa' (isim kodu) yerinde kalmış; 'V' (fiil kodu) ve ona bağlı 'uç' dizgisi silinmiştir. Böylece, eş yazımlı *'fly'* kelimesinin üç farklı anlamından geriye sadece 'sinek' dizgisi ve anlam seçiminin tamamlandığına ('NNa' kodu, anlam seçiminden sonra 'Na' [ikame kodu] ile değiştirilmiştir.) işaret eden 'Na' kodu kalmıştır. (iii) Analizinin devamı '*old*' kelimesinin hangi anlamda kullanılacağı (yaşlı/eski) ile alakalıdır. Üçüncü sıradaki kural (3.37) 'Na', 'J', 'Ja' kodlarını yan yana gördüğü yerde “eski” dizgisinin bağlı olduğu 'J' kodunu silip “yaşlı” dizgisinin bağlı olduğu 'Ja' kodunu bırakmıştır. Kalan 'Ja' kodu ise mana tercihinin yapıldığına işaret eden 'JJa' (ikame kodu) ile değiştirilmiştir. (ix) Dördüncü ve son aşamada dördüncü kural (3.38) işletilmiş ve 'yaşlı' ile 'sinek' dizgileri 'Na' kodunda birleştirilerek HD dizgisinin üretimi gerçekleştirilmiştir.

Ancak; yukarıdaki (Şekil 3.13) örnekte 'Na' kodlu bir 'sinek' dizgisi yerine aşağıdaki (Şekil 3.14) örnekte olduğu gibi 'N' kodlu ('N' kodu akılsız varlıklarda kullanılır.) bir 'masa' dizgisi olmuş olduğunu düşünelim. Bu durumda anlama bağlı kod-dizgi seçimi farklı şekilde ilerleyecektir (Şekil 3.14):



**Şekil 3.14: Tamlanana göre tamlayanın seçildiği analiz örneği**

Görselin (Şekil 3.14) sağındaki kural kodlarında da görüldüğü gibi algoritma, 'N', 'J', 'Ja' kodlarını yan yana gördüğü yerde (karar kudundaki [?X\_] “gör, yaz, sil” kuralına göre) 'Ja' kodunu silip 'J' kodunu bırakacak ve kalan 'J' kodunu ise mana tercihinin

yapıldığına işaret eden 'JJ' (ikame kodu) ile değiştirecektir. Yani 'old' kelimesi için "yaşlı" kalkmış ve "masa" için tamlayan olarak "eski" dizgisi seçilmiş olacaktır.

### 3.2.6.3. Birden fazla eş yazımlı bulunduran dizgi çevirileri:

Şekil 3.13 ile işaret edilen örnekte eş yazımlı bir kelimenin (*fly*) doğru çevirisinin başındaki sıfat (*old*) yardımıyla belirlendiği analiz süreci tarif edilmiştir. Ancak; bazı KD dizgilerinde birden fazla eş yazımlı bulunabilmekte ve her eş yazımlının, bulunduğu konumu dikkate alınarak doğru anlamıyla çevrilmesi gerekmektedir. Örneğin "*People at my work object to the book.*" (İş yerimdeki insanlar kitaba itiraz eder.) dizgisinde üç eş yazımlı bulunur. Bunlar 'work', 'object' ve 'book' kelimeleridir. Hatta bu kelimelerden ikisi (*work, object*) yan yanadır.

"*People at my work object to the book.*" dizgisinin önerilen tez modeli makine çevirisinde girdi olarak kullanıldığını düşünelim. Model, ayrıştırma ünitesine (3.1) aldığı dizgideki her bir eş yazımlı için üç farklı anlam ihtimali sıralayacaktır. Bu, her eş yazımlı kelimenin ikişer tane yanlış çevirisi olduğu anlamına gelir. Yani, üç ayrı eş yazımlının bulunduğu bir cümlede toplam altı yanlış kelime çevirisi aday olacaktır. Bu altı yanlış çeviri adayının, doğru çevirilerle karıştırılmadan dizgilerden ayıklanabilmesi için altı aşamalı anlam seçimi analizine ihtiyaç vardır.

Söz konusu anlam seçimlerinin gerçekleştirildiği analiz aşamaları Şekil 3.15 ile bu kez gerçek liste elemanları görüntüsü ile resmedilmiş ve dizgilerin bir araya gelmesinde başrol oynayan kural kodları (3.39) ise (birleştirme matrisindeki [3.2.2] gibi bir iç içe liste veri tipinde) aşağıda sunulmuştur. Tüm dizgiyi analiz eden kural listelerinin olduğu bu iç içe listede (3.39) eş yazımlıların çözümlenmesinde yararlanılan kural kodları, ayırt edilebilmesi için, kalın harf karakterleriyle gösterilmiştir.

(3.39)

['V', 'V1', 'NN', 'DT', 'X\_??', 'V', "", "", "", ""],  
 ['V', 'V1', 'NN', 'DT', 'X\_??', 'V', "", "", "", ""],  
 ['V', 'NN', 'DT', '\_X?', 'N', "", "", "", ""],  
 ['V', 'NN', 'DT', '\_X?', 'N', "", "", "", ""],  
 ['N', 'DT', "", '><', 'WT', "", "", "", ""],  
 ['WT', 'TO', 'WR', '?\_X', 'WR', "", "", "", ""],  
 ['N', 'DT', "", '><', 'WT', "", "", "", ""],  
 ['V', 'V1', 'NN', 'WT', '??\_X', 'WT', "", "", "", ""],  
 ['WT', 'WR', "", '><', 'WTWR', "", "", "", ""],  
 ['WTWR', 'V', 'NN', '?X\_', 'V', "", "", "", ""],  
 ['WT', 'PRP', "", '><', 'WTPRP', "", "", "", ""],  
 ['WTPRP', 'Na', "", '><', 'WO', "", "", '--ki', ""],  
 ['WTWR', 'V', "", '><', 'VO', "", "", "", ""],  
 ['VO', 'WO', "", ':', 'SVO', "", "", '--er'] ]

[ 'book', 'book', 'book', 'the', 'to', 'to', 'object', 'object', 'object', 'work', 'work', 'work', 'my', 'at', 'people', '[]'  
 [ 'V', 'V1', 'NN', 'DT', 'TO', 'WR', 'V', 'V1', 'NN', 'V', 'V1', 'NN', 'DT', 'PRP', 'Na', '[]'  
 [ 'kaydet', 'rezerve et', 'kitap', 'işbu', ", '\_a', 'itiraz et', 'reddet', 'eşya', 'çalış', 'çalıştır', 'iş', 'benim', '\_da', 'insanlar', '[]'

[ 'V', **1** 'NN', 'DT', 'TO', 'WR', 'V', 'V1', 'NN', 'V', 'V1', 'NN', 'DT', 'PRP', 'Na', '[]'  
 [ 'kaydet', 'kitap', 'işbu', ", '\_a', 'itiraz et', 'reddet', 'eşya', 'çalış', 'çalıştır', 'iş', 'benim', '\_da', 'insanlar', '[]'

[ 'V', 'NN', 'DT', 'TO', 'WR', 'V', 'V1', 'NN', 'V', **2** 'NN', 'DT', 'PRP', 'Na', '[]'  
 [ 'kaydet', 'kitap', 'işbu', ", '\_a', 'itiraz et', 'reddet', 'eşya', 'çalış', 'iş', 'benim', '\_da', 'insanlar', '[]'

**3** [ 'N', 'DT', 'TO', 'WR', 'V', 'V1', 'NN', 'V', 'NN', 'DT', 'PRP', 'Na', '[]'  
 [ 'kitap', 'işbu', ", '\_a', 'itiraz et', 'reddet', 'eşya', 'çalış', 'iş', 'benim', '\_da', 'insanlar', '[]'

[ 'N', 'DT', 'TO', 'WR', 'V', 'V1', 'NN', **4** 'N', 'DT', 'PRP', 'Na', '[]'  
 [ 'kitap', 'işbu', ", '\_a', 'itiraz et', 'reddet', 'eşya', 'iş', 'benim', '\_da', 'insanlar', '[]'

[ 'WT', 'TO', 'WR', 'V', 'V1', 'NN', 'N', 'DT', 'PRP', 'Na', '[]'  
 [ 'kitap', ", '\_a', 'itiraz et', 'reddet', 'eşya', 'iş', 'benim', '\_da', 'insanlar', '[]'

[ 'WT', 'WR', 'V', 'V1', 'NN', 'N', 'DT', 'PRP', 'Na', '[]'  
 [ 'kitap', '\_a', 'itiraz et', 'reddet', 'eşya', 'iş', 'benim', '\_da', 'insanlar', '[]'

[ 'WT', 'WR', 'V', 'V1', 'NN', 'WT', 'PRP', 'Na', '[]'  
 [ 'kitap', '\_a', 'itiraz et', 'reddet', 'eşya', 'işim', '\_da', 'insanlar', '[]'

[ 'WT', 'WR', 'V', 'V1', **5** 'WT', 'PRP', 'Na', '[]'  
 [ 'kitap', '\_a', 'itiraz et', 'reddet', 'işim', '\_da', 'insanlar', '[]'

[ 'WTWR', 'V', 'V1', 'WT', 'PRP', 'Na', '[]'  
 [ 'kitaba', 'itiraz et', 'reddet', 'işim', '\_da', 'insanlar', '[]'

[ 'WTWR', 'V', **6** 'WT', 'PRP', 'Na', '[]'  
 [ 'kitaba', 'itiraz et', 'işim', '\_da', 'insanlar', '[]'

[ 'WTWR', 'V', 'WTPRP', 'Na', '[]'  
 [ 'kitaba', 'itiraz et', 'işimde', 'insanlar', '[]'

[ 'WTWR', 'V', 'WO', '[]'  
 [ 'kitaba', 'itiraz et', 'işimdeki insanlar', '[]'

[ 'VO', 'WO', '[]'  
 [ 'kitaba itiraz et', 'işimdeki insanlar', '[]'

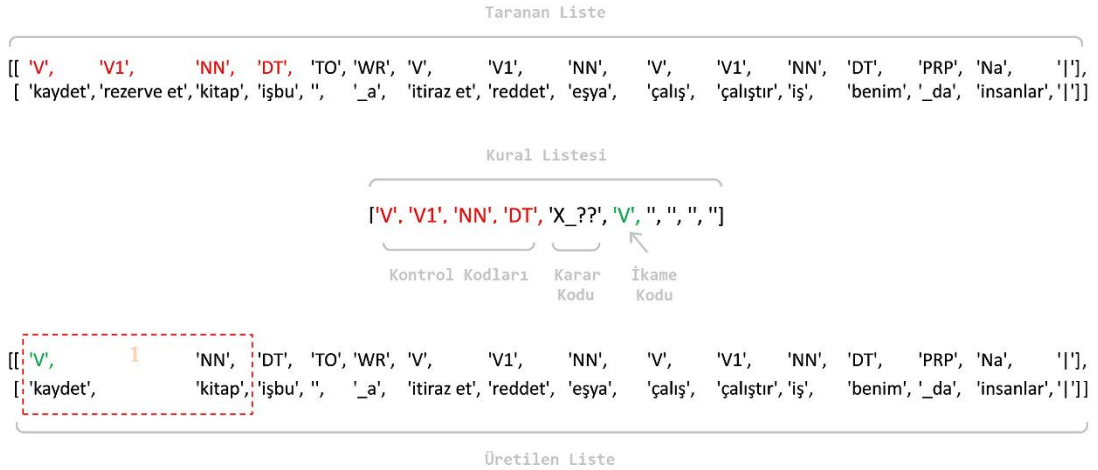
[ 'SVO', '[]'  
 [ 'işimdeki insanlar kitaba itiraz eder', '[]'

Şekil 3.15: Üç farklı eş yazımlı için doğru anlam seçimi analiz aşamaları

Altı ayrı kademedeki gerekleŒen eŒ yazımlı analizi aŒamaları yukarıdaki Œekil 3.15 iinde (numaralanmıŒ) kutularla iŒaretlenmiŒtir. İŒaretli alanlarda “book”, “object” ve “work” eŒ yazımlı kelimelerinin dođru anlam seimleri yapılmaktadır. Œekil 3.15’deki bu analiz sreleri altı farklı grsel eŒliđinde aŒađıda tarif edilmiŒtir.

Altı adet grselin her birinde kod ve dizgi listeleri; bu listelerin analizinde kullanılan kural listesi (Kural listeleri iin ‘3.2.2. BirleŒtirme matrisi’ incelenebilir.) ve yapılan analiz sonucunda ortaya ıkan yeni kod ve dizgi listeleri sıralanır. Œekiller tarif edilirken her bir grseldeki birinci i ie liste iin “taranan liste”; ortadaki kural listesinden sonraki i ie liste iin de “retilen liste” ifadesi tercih edilmiŒtir. Analiz srelerinin daha anlaşılır olması iin dizgi retimi aŒamalarında dilbilimsel bilgiye mstenit tanımlamalara yer verilmiŒ ve retim iŒlemlerinin kodlar zerinden nasıl gerekleŒtiđi tarif edilmiŒtir.

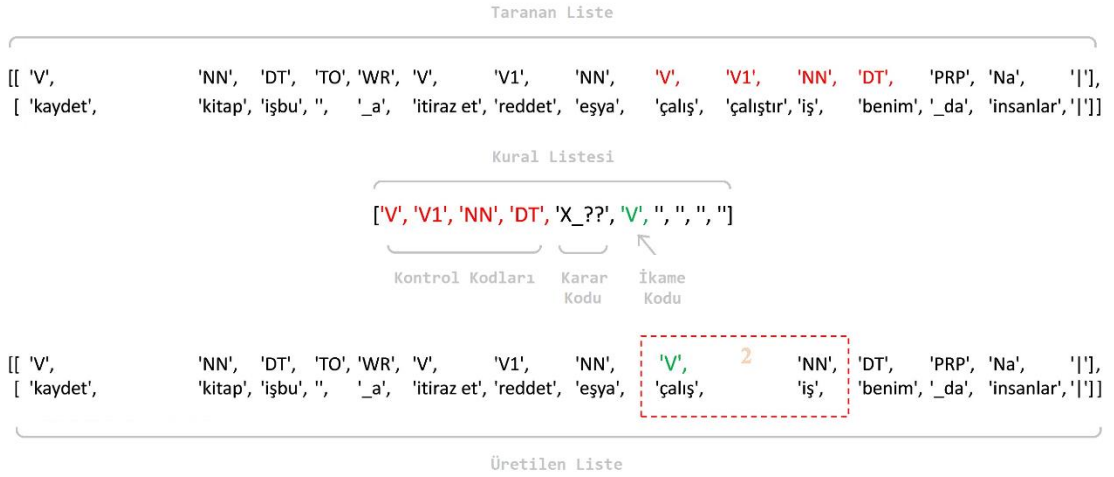
AŒađıdaki (Œekil 3.16) birinci eŒ yazımlı analizinde “book” kelimesinden  eviri adayları ('kitap', 'rezerve et', 'kaydet') ortaya ıkmıŒtır. Ancak; bu adayların hemen nndeki (yazım dilinde “the” kelimesinin karŒılıđı olan) “iŒbu” artikeli, fiil ifadelerinin nnde olamayacađı iin 'rezerve et' ve 'kaydet' fiillerinin dođru eviri olma Œansı yoktur. Bu fiiller listeden sırayla kaldırılmalıdır. Bu nedenle taranan listesinde grlen 'rezerve et' dizgisi ve bađlı olduđu kod (V1) retilen listede silinmiŒtir (Œekil 3.16).



**Şekil 3.16: “kitap”, “rezerve et” “kaydet” çeviri adaylarının önünde “işbu” artikeli bulunduğu için üç çeviri ihtimalinden ikincisi olan “rezerve et” dizgisinin silinmesi**

Sistemde kodlar üzerinden birinci analiz süreci şu şekildedir: Yukarıdaki (Şekil 3.16) taranan listede dört dizgi kodunun ('V', 'V1', 'NN', 'DT') kural listesinde bulunan kontrol kodlarıyla ('V', 'V1', 'NN', 'DT') eşleştiği görülür. Eşleşen kontrol kodlarının yanında hangi karar kodu (X\_??) varsa, taranan listedeki kodlar bu karar koduna göre düzenlenmelidir. Çünkü, taranan listedeki dört kodun, yan yana görülmesi durumunda üretilen listeye nasıl geçirilmesi gerektiği karar kodunda belirlenmiştir. Mevcut karar kodunda “yaz, sil, gör, gör” komut karakterleri bulunduğu için taranan listedeki 'V', üretilen listeye yazılmış; taranan listedeki 'V1', üretilen listeden silinmiş ve taranan listede 'NN' ile 'DT'nin mevcut olduğu görülmüştür.

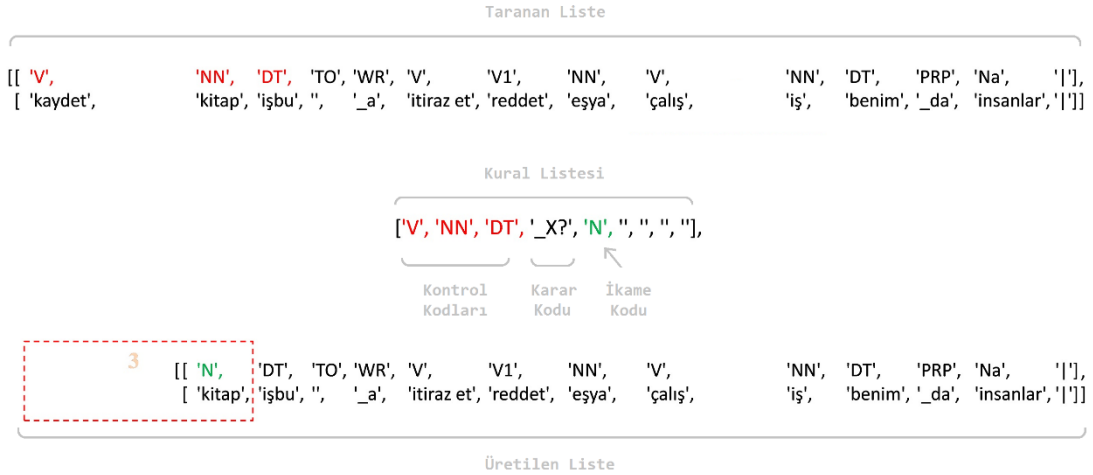
Aşağıdaki (Şekil 3.17) ikinci eş yazımlı analizinde “work” kelimesinden üç çeviri adayı ('iş', 'çalıştır', 'çalış') çıkmıştır. Yalnız; bu adayların hemen önünde konumlanan (yazım dilinde “the” kelimesinin karşılığı olan) “işbu” artikelinin, fiil ifadelerinin önünde kullanılamayacağı için 'çalıştır' ve 'çalış' fiillerinin doğru çeviri olma ihtimali ortadan kalkmıştır. Bu fiillerin listeden sırayla kaldırılması gerekir. Bu nedenle tarama listesinde bulunan 'çalıştır' dizgisi ve bağlı olduğu kod (V1) üretilen listede silinmiştir (Şekil 3.17).



**Şekil 3.17: “iş”, “çalıştır” “çalış” çeviri adaylarının önünde “benim” sahiplik sıfatı bulunduğu için üç çeviri ihtimalinden ikincisi olan “çalıştır” dizgisinin silinmesi**

Sistemde kodlar üzerinden ikinci analiz süreci şu şekildedir: Yukarıdaki (Şekil 3.17) taranan listede dört dizgi kodu ('V', 'V1', 'NN', 'DT'), kural listesinde bulunan kontrol kodlarıyla ('V', 'V1', 'NN', 'DT') eşleşmiştir. Eşleşen kontrol kodlarının yanında hangi karar kodu (X\_??) varsa, taranan listedeki kodlar bu karar koduna göre tanzim edilmelidir. Çünkü, taranan listedeki dört kodun, yan yana görülmesi durumunda üretilen listeye nasıl geçirilmesi gerektiği karar kodunda belirlenmiştir. Mevcut karar kodunda “yaz, sil, gör, gör” komutları olduğu için taranan listedeki 'V', üretilen listeye yazılmış; taranan listedeki 'V1', üretilen listeden silinmiş ve taranan listede 'NN' ile 'DT'nin mevcut olduğu görülmüştür.

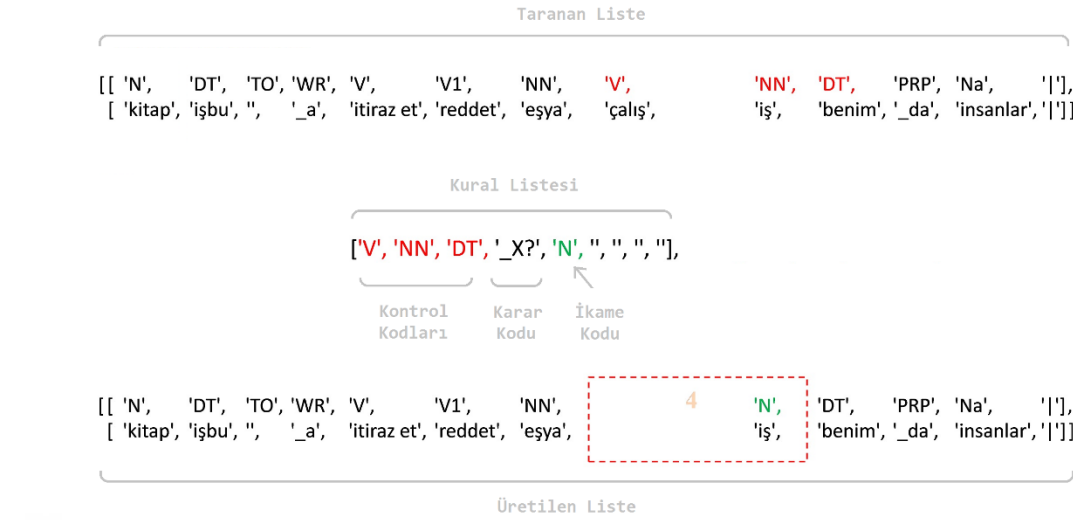
Aşağıdaki (Şekil 3.18) üçüncü eş yazımlı analizinde “book” kelimesi için tarama listesinde iki çeviri adayı ('kitap', 'kaydet') görülmektedir. Ancak; “işbu” artikelinin bir fiilden önce görülmesi mümkün olmadığı için 'kaydet' fiilinin doğru çeviri olma ihtimali ortadan kalkmış ve tarama listesinde bulunan bu dizgi ve bağlı olduğu kod (V) üretilen listede silinmiştir (Şekil 3.18).



**Şekil 3.18: “kitap” ve “kaydet” çeviri adaylarının önünde “işbu” artikeli bulunduğu için iki çeviri ihtimalinden ikincisi olan “kaydet” dizgisinin silinmesi**

Sistemde kodlar üzerinden üçüncü analiz süreci şu şekilde gerçekleşmiştir: Yukarıdaki (Şekil 3.18) taranan listede üç dizgi kodu ('V', 'NN', 'DT'), kural listesinde bulunan kontrol kodlarıyla ('V', 'NN', 'DT') eşleşmiştir. Taranan listedeki üç kodun yan yana görüldüğünde, üretilen listeye nasıl geçirilmesi gerektiği karar kodunda (\_X?) belirlenmiştir. Karar kodunda “sil, yaz, gör” komutları olduğu için taranan listedeki 'V', üretilen listeden silinmiş; taranan listedeki 'NN', üretilen listeye 'N' (ikame kodu) ile yazılmış ve 'DT' nin taranan listede mevcut olduğu görülmüştür.

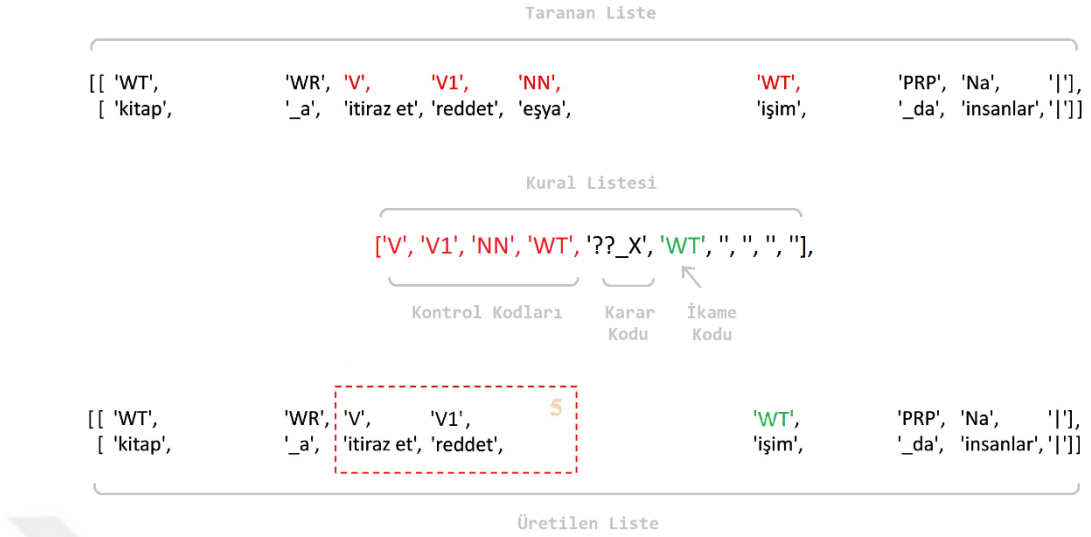
Aşağıdaki (Şekil 3.19) dördüncü eş yazımlı analizinde “work” kelimesi için tarama listesinde iki çeviri adayı ('iş', 'çalış') görülmektedir. Ancak; “benim” sahiplik sıfatının fiilden önce bulunması mümkün olmadığı için 'çalış' dizgisinin doğru olma ihtimali ortadan kalkmış ve tarama listesinde bulunan bu dizgi ve bağlı olduğu kod (V) üretilen listede silinmiştir (Şekil 3.19).



**Şekil 3.19: “iş” ve “çalış” çeviri adaylarının önünde “benim” sahiplik sıfatı bulunduğu için iki çeviri ihtimalinden ikincisi olan “çalış” dizgisinin silinmesi**

Sistemde kodlar üzerinden dördüncü analiz süreci şu şekildedir: Yukarıdaki (Şekil 3.19) taranan listede üç dizgi kodu ('V', 'NN', 'DT'), kural listesindeki kontrol kodlarıyla ('V', 'NN', 'DT') eşleşmiştir. Taranan listedeki üç kodun yan yana görüldüğünde, üretilen listeye nasıl geçirilmesi gerektiği karar kodunda (\_X?) belirlenmiştir. Karar kodu “sil, yaz, gör” komut karakterlerini içerdiği için taranan listedeki 'V', üretilen listeden silinmiş; taranan listedeki 'NN', üretilen listeye 'N' (ikame kodu) ile yazılmış ve 'DT' nin taranan listede mevcut olduğu görülmüştür.

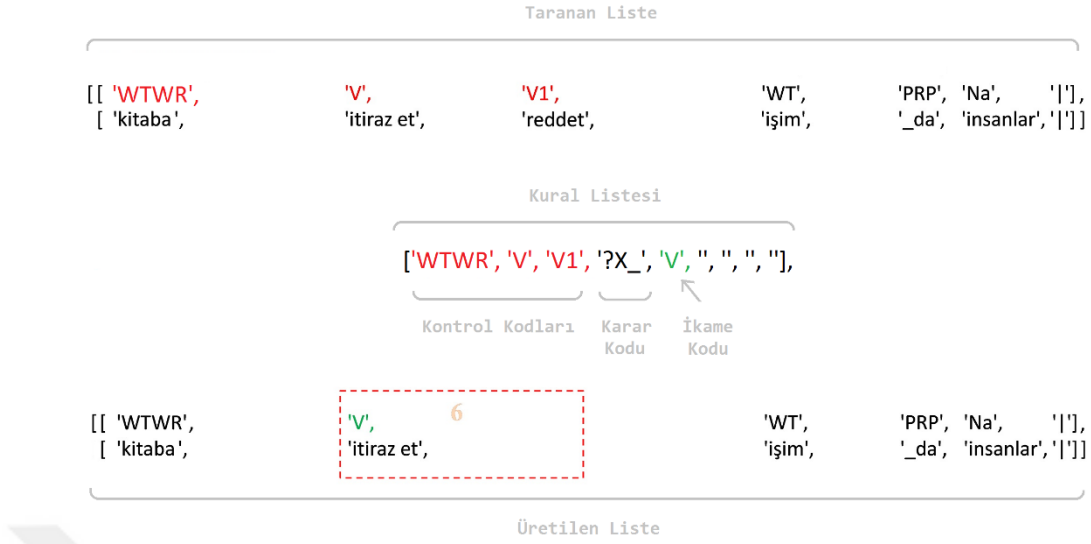
Aşağıdaki (Şekil 3.20) beşinci eş yazımlı analizinde “*object*” kelimesi için tarama listesinde üç çeviri adayı ('eşya', 'reddet', 'itiraz et') görülmektedir. Ancak; “işim” kelimesi gibi sahiplik eki almış (özne adayı) cins isimden sonra eş yazımlının yalnız fiil anlamları ('reddet', 'itiraz et') kullanılabileceği için 'eşya' dizgisinin doğru olma ihtimali ortadan kalkmış ve tarama listesinde bu dizgi ve bağlı olduğu kod (NN) üretilen listede silinmiştir (Şekil 3.20).



**Şekil 3.20: “eşya”, “reddet” “itiraz et” çeviri adaylarının önünde “işim” cins ismi bulunduğu için üç çeviri ihtimalinden birincisi olan “eşya” dizgisinin silinmesi**

Sistemde kodlar üzerinden beşinci analiz süreci şu şekildedir: Yukarıdaki (Şekil 3.20) taranan listede dört dizgi kodu ('V', 'V1', 'NN', 'WT'), kural listesindeki kontrol kodlarıyla ('V', 'V1', 'NN', 'WT') eşleşmiştir. Taranan listedeki dört kodun yan yana görüldüğünde, üretilen listeye nasıl geçirilmesi gerektiği karar kodunda (??\_X) belirlenmiştir. Karar kodu “gör, gör, sil, yaz” komutlarını içerdiği için taranan listede 'V' ve 'V1'in olduğu görülmüş; taranan listedeki 'NN' üretilen listede silinmiş ve taranan listedeki 'WT', üretilen listeye yine 'WT' (ikame kodu) ile yazılmıştır.

Aşağıdaki (Şekil 3.21) altıncı eş yazımlı analizinde “*object*” kelimesi için tarama listesinde iki çeviri adayı ('itiraz et', 'reddet') görülmektedir. Ancak; bu iki adaydan sonra gelen 'kitaba' dizgisi ismin yönelme hâlidir ve nesne olarak kullanılamaz. Yani, geçişli fiil olan ve nesne alması gereken 'reddet' dizgisinin çeviri ihtimali olma durumu ortadan kalkmış ve tarama listesinde bulunan bu dizgi ve bağlı olduğu kod (V1) üretilen listede silinmiştir (Şekil 3.21).



**Şekil 3.21: “itiraz et” ve “reddet” çeviri adaylarının ardında yönelme durumu bildiren “kitaba” dizgisi olduğu için (yalnız belirtme hali ile kullanılabilen) “reddet” dizgisinin silinmesi**

Sistemde kodlar üzerinden altıncı analiz süreci şu şekilde gelişmiştir: Yukarıdaki (Şekil 3.21) taranan listede üç dizgi kodu ('WTWR', 'V', 'V1'), kural listesindeki kontrol kodlarıyla ('WTWR', 'V', 'V1') eşleşmiştir. Taranan listedeki üç kodun yan yana görüldüğünde, üretilen listeye nasıl geçirilmesi gerektiği karar kodunda (?X\_) belirlenmiştir. Karar kodunda “gör, yaz, sil” komut karakterleri bulunduğu için taranan listede 'WTWR'nin mevcut olduğu görülmüş; taranan listedeki 'V', üretilen listeye yine 'V' (ikame kodu) ile yazılmış ve 'V1' taranan listede silinmiştir.

## **DÖRDÜNCÜ BÖLÜM**

### **UYGULAMA**

Bu tez çalışmasında makine çevirilerinde sorun teşkil eden eş yazımlı bulunduran KD ifadelerinin kural tabanlı bir model yaklaşımı ile doğru şekilde hedef dile çevrilebilmesinin mümkün olup olmadığı sorusu araştırılmaktadır. Bu amaçla tez çalışmasında önerilen kural tabanlı bir makine çevirisi modeli tasarlanmıştır. Tasarlanan Kural Matrisli Makine Çevirisi (KMMÇ) modeli ile eş yazımlı kelimeler bulunduran cümle çevirileri yapılarak aynı cümlelerin Google, Bing ve Yandex çevirileri ile karşılaştırılması sağlanmıştır. Karşılaştırma için en iyi değerlendirme yöntemi olarak kabul edilen çevirmen değerlendirmesinden yararlanılmıştır.

Tezin “Uygulama” bölümü ‘KMMÇ Modelinin Test Aşaması ve Kullanılan Yöntem’ (4.1), ‘Çeviri Girdileri İçin Eş Yazımlı Kelimeler ve Örnek Cümleler’ (4.2), ‘Çeviri Çıktıları ve Çevirmen Değerlendirme Ölçeği’ (4.3), ‘Çevirmen Değerlendirme Puanları’ (4.4) ile ‘Eş Yazımlı Kelime Bulunduran Çeviri Hatalarının İncelenmesi’ (4.5) alt başlıklarıyla ele alınmıştır.

#### **4.1. Önerilen KMMÇ Modelinin Test Aşaması ve Kullanılan Yöntem**

Çevirmenlere göre bir metnin iki ayrı kişi tarafından aynı şekilde çevrilebilmesi mümkün değildir. Bu nedenle çeviri değerlendirmelerinde öznel yaklaşımların önüne geçebilmek mümkün görünmemektedir. Makine Çevirileri için bu durum daha da karmaşık bir hal arz etmektedir. Öyle ki; bir makine çevirisi çıktısının doğru çevrildiğine yalnız kaynak ve hedef dili çok iyi bilen kişiler karar verebilir. Bir başka deyişle, makine çevirisi çıktıları için "kesin referans" olarak nitelendirilebilecek, üzerinde anlaşılmiş uluslararası standartların varlığından söz edilemez. Bu nedenle hangi makine çevirisinin diğerinden daha iyi olduğuyla alakalı öznel yargılara alışık olmakta yarar vardır. Burada hedef, MÇ ölçümlerinde insana ait dilbilimsel sezgilerden yeterince yararlanabilmenin yolunu bulmaktır (White, 2003: 211 - 242).

Makine çevirilerinde girdi ve çıktı arasındaki ilişkinin değerlendirilme şekli, ölçümün kara kutu (*black box*) testi ya da cam kutu (*glass box*) testi olup olmadığını belirleyen faktördür. Kara kutu incelemelerinde, çeviri motorunun mekaniği hesaba katılmadan yalnız girdi ve çıktıların değerlendirilmesi yapılırken; cam kutu incelemelerinde MÇ modelinin içine bakılarak her bir bileşenin çeviri sürecinde kendisinden bekleneni yapıp yapmadığı ile ilgilenilir. Her iki yöntemin farklı avantajları vardır. Cam kutu araştırmaları daha çok sistemlerin kapsamının genişletilebilirliği gibi durumların belirlenmesinde kullanılır. Kara kutu testleri, araştırılan sistemin yapısını veya tasarım felsefesini dışarıda tuttuğu için farklı sistemlerin karşılaştırılması veya belirli bir modelin çeviri başarısının tespiti için daha uygundur (A.g.e.). Dolayısıyla bu tez çalışmasında önerilen KMMC modelinin başarı ölçümleri için, karşılaştırıldığı diğer modellerden farklı olarak kural tabanlı bir yapı arz etmesi ve birbirinden uzak dil yapılarına sahip dillerin çevirisi için tasarlanmış olması nedeniyle, bir makro değerlendirme yaklaşımı olan kara kutu testi uygun görülmüştür. Testin uygulanabilmesi için öncelikle Google, Bing Yandex ve KMMC modellerinin çeviri başarılarının karşılaştırılmasında kullanılacak olan eş yazımlı kelimeler bulunduran cümlelerin tespiti gerçekleştirilmiştir.

#### **4.2. Çeviri Girdileri İçin Eş Yazımlı Kelimeler ve Örnek Cümleler**

Eş yazımlı kelimeler ve bu kelimelerin kullanıldığı örnek cümlelerin tespitinde dünya çapında bilinen sözlüklerin (Oxford, 2021; Cambridge, 2021; Macmillan, 2021; Merriam-Webster, 2021; Collins, 2021; Yourdictionary, 2021) çevrimiçi sürümlerinden yararlanılmış ve cümle tercihlerinde kısa ve karmaşık olmayan örneklerin seçilmesine özen gösterilmiştir. Bazı cümlelerde özne veya nesne gruplarında eklemeler yapılmış; ancak eklemeler yapılırken cümlelerin sözlük örneklerindeki yapısal özelliklerine sadık kalınmıştır. Toplam 80 eş yazımlı kelime bulunduran cümle bir araya getirilmiş ve eş yazımlı kelimelerin, isim ifadeleri (*noun phrase*) içerisinde kullanılmış olanları Tablo 4.1; fiil ifadeleri (*verb phrase*) içerisinde kullanılmış olanları ise Tablo 4.2 ile aşağıda sunulmuştur.

**Tablo 4.1: Eş yazımlı kelimeler ve isim ifadeleri içinde kullanıldıkları cümleler**

EŞ YAZIMLI KELİMELER VE ANLAMLARI	EŞ YAZIMLI KELİMELERİN BULUNDUĞU ÖRNEK CÜMLELER
1 <b>arms:</b> <i>kollar(i)</i> , donatmak( <i>f</i> )	She was holding the baby in her <b>arms</b> .
2 <b>bear:</b> <i>ayı(i)</i> , katlanmak( <i>f</i> )	We saw the little <b>bear</b> cub.
3 <b>beef:</b> <i>et(i)</i> , yakınmak( <i>f</i> )	My friend is a <b>beef</b> producer.
4 <b>boil:</b> <i>çıban(i)</i> , kaynamak( <i>f</i> )	The <b>boil</b> on his arm was painful.
5 <b>book:</b> <i>kitap(i)</i> , rezerve etmek( <i>f</i> )	He is reading a <b>book</b> about American history.
6 <b>box:</b> <i>kutu(i)</i> , boks yapmak( <i>f</i> )	The mailman kept the letter in a <b>box</b> in his room.
7 <b>carp:</b> <i>sazan(i)</i> , mızımızlanmak( <i>f</i> )	<b>Carp</b> is a fish.
8 <b>clip:</b> <i>kesit(i)</i> , <b>klip(i)</b> , kesmek( <i>f</i> )	A <b>clip</b> from that video.
9 <b>crane:</b> <i>turna(i)</i> , boynunu uzatmak( <i>f</i> )	That bird is a <b>crane</b> .
10 <b>express:</b> <i>ekspres(i)</i> , <i>hızlı(s)</i> , ifade etmek( <i>f</i> )	An <b>express</b> shipment.
11 <b>drop:</b> <i>damla(i)</i> , düşürmek( <i>f</i> )	The last <b>drop</b> of my coffee.
12 <b>fast:</b> <i>hızlı(s)</i> , oruç tutmak( <i>f</i> )	My friends at work love <b>fast</b> cars.
13 <b>flag:</b> <i>bayrak(i)</i> , işaret etmek( <i>f</i> )	A green <b>flag</b> indicates the start of the race.
14 <b>kid:</b> <i>çocuk(i)</i> , kandırmak( <i>f</i> )	A <b>kid</b> in high school.
15 <b>leaves:</b> <i>yapraklar (i)</i> , ayrılmak( <i>f</i> )	The children love to play in the <b>leaves</b> .
16 <b>mean:</b> <i>kaba(s)</i> , anlamına gelmek( <i>f</i> )	She was very <b>mean</b> to her sister.
17 <b>move:</b> <i>hareket(i)</i> , hareket ettirmek( <i>f</i> ), taşımak( <i>f</i> )	Every <b>move</b> was painful for the old man.
18 <b>object:</b> <i>nesne(i)</i> , itiraz etmek( <i>f</i> )	A collection of precious <b>objects</b> .
19 <b>page:</b> <i>sayfa(i)</i> , çağrı yapmak( <i>f</i> )	Fill in the form on the <b>page</b> .
20 <b>pine:</b> <i>çam(i)</i> , yas tutmak( <i>f</i> )	A <b>pine</b> forest in the city.
21 <b>press:</b> <i>basın(i)</i> , <b>basma(i)</b> , basmak( <i>f</i> )	At the <b>press</b> of a button.
22 <b>desert:</b> <i>çöl(i)</i> , terk etmek( <i>f</i> )	The burning <b>desert</b> sun.
23 <b>ships:</b> <i>gemi(i)</i> , göndermek ( <i>f</i> )	Build a <b>ship</b> for the navy.
24 <b>tie:</b> <i>kravat(i)</i> , <i>bağ(i)</i> , bağlamak( <i>f</i> )	I hate wearing a pink <b>tie</b> .
25 <b>tie:</b> <i>kravat(i)</i> , <b>bağ(i)</b> , bağlamak( <i>f</i> )	They don't know the <b>tie</b> between my grandmother and the town.
26 <b>buffet:</b> <i>büfe(i)</i> , vurmak( <i>f</i> )	Dinner will be a cold <b>buffet</b> .
27 <b>close:</b> <i>yakın(s)</i> , kapatmak( <i>f</i> )	My uncle and my father are very <b>close</b> .
28 <b>console:</b> <i>konsol(i)</i> , avunmak( <i>f</i> ), avutmak( <i>f</i> )	The antique <b>console</b> will be in the living room.
29 <b>record:</b> <b>rekor(i)</b> , kaydetmek( <i>f</i> )	This is the new word <b>record</b> .
30 <b>fly:</b> <i>sinek(i)</i> , uçmak( <i>f</i> ), uçurmak( <i>f</i> )	I swatted the <b>fly</b> with a newspaper.
31 <b>wind:</b> <i>rüzgar(i)</i> , sarmak( <i>f</i> )	The leaves rustle in the <b>wind</b> .
32 <b>does:</b> <b>Geyikler(i)</b> , yapar( <i>f</i> )	The <b>does</b> are walking in the forest.
33 <b>park:</b> <i>park(i)</i> , park etmek( <i>f</i> ), bırakmak( <i>f</i> ), yerleşmek( <i>f</i> )	They sat on the grass in the <b>park</b> .
34 <b>house:</b> <i>ev(i)</i> , barındırmak ( <i>f</i> )	The gas stove was in the old <b>house</b> in the village.
35 <b>live:</b> <i>naklen(s)</i> , <b>yaayında(z)</b> , yaşamak( <i>f</i> )	They will be <b>live</b> on the city radio.
36 <b>type:</b> <i>tip(i)</i> , yazmak( <i>f</i> )	That person is really not your <b>type</b> .
37 <b>jam:</b> <i>reçel(i)</i> , sıkışmak( <i>f</i> ), tutukluk yapmak( <i>f</i> )	My roommates ate two <b>jars</b> of jam in two days.
38 <b>attribute:</b> <i>özellik(i)</i> , atfetmek( <i>f</i> )	Endurance was the most important <b>attribute</b> in a teacher.
39 <b>watch:</b> <i>saat(i)</i> , seyretmek( <i>f</i> )	The old man's old <b>watch</b> is fast.
40 <b>wave:</b> <i>dalga(i)</i> , el sallamak( <i>f</i> )	There is a huge <b>wave</b> of divorce in the developed countries.

**Tablo 4.2: Eş yazımlı kelimeler ve fiil ifadeleri içinde kullanıldıkları cümleler**

EŞ YAZIMLI KELİMELER VE ANLAMLARI	EŞ YAZIMLI KELİMELERİN BULUNDUĞU ÖRNEK CÜMLELER
41 <i>arms</i> : kollar( <i>i</i> ), donatmak( <i>f</i> )	She <b>arms</b> herself with knowledge.
42 <i>bear</i> : ayı( <i>i</i> ), katlanmak( <i>f</i> )	They <b>bear</b> the pain.
43 <i>beef</i> : et( <i>i</i> ), yakınmak( <i>f</i> )	They are going to <b>beef</b> .
44 <i>boil</i> : çıban( <i>i</i> ), kaynamak ( <i>f</i> )	They <b>boil</b> the goat milk in a pan.
45 <i>book</i> : kitap( <i>i</i> ), rezerve etmek( <i>f</i> )	We <b>book</b> flights to big cities.
46 <i>box</i> : kutu( <i>i</i> ), boks yapmak( <i>f</i> )	He is going to <b>box</b> in the Olympics in summer.
47 <i>carp</i> : sazan( <i>i</i> ), mızızlanmak( <i>f</i> )	The people <b>carp</b> about junk mail.
48 <i>clip</i> : kesit, klip( <i>i</i> ), kesmek( <i>f</i> )	Her mom is going to <b>clip</b> her hair.
49 <i>crane</i> : turna( <i>i</i> ), boynunu uzatmak( <i>f</i> )	They <b>crane</b> forward.
50 <i>express</i> : ekspres( <i>i</i> ), ifade etmek( <i>f</i> )	I'd like to <b>express</b> my gratitude to everyone.
51 <i>drop</i> : damla ( <i>i</i> ), düşürmek( <i>f</i> )	She is going to <b>drop</b> the book.
52 <i>fast</i> : hızlı( <i>s</i> ), oruç tutmak( <i>f</i> )	Muslims <b>fast</b> during Ramadan.
53 <i>flag</i> : bayrak( <i>i</i> ), işaretlemek ( <i>f</i> )	<b>Flag</b> the potential problems in our proposal.
54 <i>kid</i> : çocuk( <i>i</i> ), kandırmak ( <i>f</i> )	Don't <b>kid</b> yourself.
55 <i>leaves</i> : yapraklar( <i>i</i> ), ayrılmak( <i>f</i> )	Their father <b>leaves</b> home for work.
56 <i>mean</i> : kaba( <i>s</i> ), anlamına gelmek( <i>f</i> )	I didn't <b>mean</b> any harm.
57 <i>move</i> : hareket( <i>i</i> ), hareket ettirmek( <i>f</i> ), taşımak( <i>f</i> )	They <b>move</b> the flag up and down.
58 <i>object</i> : nesne( <i>i</i> ), itiraz etmek( <i>f</i> )	People at my work <b>object</b> to the book.
59 <i>page</i> : sayfa( <i>i</i> ), çağrı yapmak( <i>f</i> )	<b>Page</b> the man in the hospital.
60 <i>pine</i> : çam( <i>i</i> ), yas tutmak( <i>f</i> )	They still <b>pine</b> for their lost wealth.
61 <i>press</i> : basın( <i>i</i> ), basma( <i>i</i> ), basmak( <i>f</i> )	<b>Press</b> the button for the third floor.
62 <i>desert</i> : çöl( <i>i</i> ), terk etmek( <i>f</i> )	He <b>deserted</b> his wife and children for another woman.
63 <i>ships</i> : gemi( <i>i</i> ), göndermek( <i>f</i> )	The company <b>ships</b> its goods to twenty countries.
64 <i>tie</i> : kravat( <i>i</i> ), bağ( <i>i</i> ), bağlamak( <i>f</i> )	The robbers were going to <b>tie</b> the men to a pipe with cable.
65 <i>house</i> : ev( <i>i</i> ), barındırmak( <i>f</i> )	The buildings <b>house</b> the young students.
66 <i>buffet</i> : büfe( <i>i</i> ), vurmak( <i>f</i> ), dövmek( <i>f</i> )	The waves are going to <b>buffet</b> the shore.
67 <i>close</i> : yakın( <i>s</i> ), kapatmak( <i>f</i> )	We <b>close</b> the office at noon on Fridays.
68 <i>console</i> : konsol( <i>i</i> ), avunmak( <i>f</i> ), avutmak( <i>f</i> )	She is going to <b>console</b> herself with the cat.
69 <i>record</i> : rekor( <i>i</i> ), kaydetmek ( <i>f</i> )	These cameras <b>record</b> 25 frames per second.
70 <i>fly</i> : sinek( <i>i</i> ), uçmak( <i>f</i> ), uçurmak( <i>f</i> )	He was the first person to <b>fly</b> .
71 <i>fly</i> : sinek( <i>i</i> ), uçmak( <i>f</i> ), uçurmak( <i>f</i> )	The little girls were going to <b>fly</b> the kite.
72 <i>does</i> : Geyikler( <i>i</i> ), yapar( <i>f</i> )	My daughter <b>does</b> her homework every night.
73 <i>park</i> : park( <i>i</i> ), park etmek( <i>f</i> ), bırakmak( <i>f</i> ),	They <b>park</b> their cars in the building.
74 <i>park</i> : park( <i>i</i> ), park etmek( <i>f</i> ), bırakmak( <i>f</i> ), yerleşmek( <i>f</i> )	We are going to <b>park</b> in front of the TV.
75 <i>live</i> : naklen( <i>s</i> ), yayında( <i>z</i> ), yaşamak( <i>f</i> )	He wants to <b>live</b> the last years of his life with his dogs.
76 <i>type</i> : tip( <i>i</i> ), yazmak( <i>f</i> )	I <b>type</b> over 100 words per minute.
77 <i>jam</i> : reçel( <i>i</i> ), sıkışmak( <i>f</i> ), tutukluk yapmak( <i>f</i> )	Their man's gun will <b>jam</b> .
78 <i>attribute</i> : özellik( <i>i</i> ), atfetmek( <i>f</i> )	She will <b>attribute</b> her success to her late father's hard work.
79 <i>watch</i> : saat( <i>i</i> ), seyretmek( <i>f</i> )	Our kids only <b>watch</b> YouTube videos at weekends.
80 <i>wave</i> : dalga( <i>i</i> ), el sallamak( <i>f</i> )	The people on the bus <b>wave</b> to the villagers.

Yukarıdaki tablolarda (Tablo 4.1, Tablo 4.2) numaralandırılmış şekilde sıralanan eş yazımlı kelimeler kalın ve yatık olarak gösterildikleri gibi, örnek cümlelerde de kalın olarak yazılmıştır. Eş yazımlı bir kelimenin her iki tabloda kullanıldığı durumlarda cümlelerin, tablolar arası karşılaştırılabilmesi için listelerde aynı sıralamada olmalarına özen gösterilmiştir.

Oluşturulan bu listeler (Tablo 4.1, Tablo 4.2) Google, Bing, Yandex ve KMMÇ sistemlerinde KD girdisi olarak kullanılmış; elde edilen HD çeviri çıktılarının bir araya getirilmesiyle çevirmen derecelendirme ölçütüne (Tablo 4.3) dayanan bir çeviri değerlendirme ölçeği (Ek: 1- Ek: 10) hazırlanmıştır.

### **4.3. Çeviri Çıktıları ve Çevirmen Değerlendirme Ölçeği**

Eş yazımlı kelimeler bulunduran 40 adet isim grubu (*noun phrase*), 40 adet fiil grubu (*verb phrase*) içeren cümle girdileri (Tablo 4.1, Tablo 4.2) tez çalışmasında önerilen KMMÇ modeliyle İngilizceden Türkçeye çevrilmiştir. Aynı girdilerin (Tablo 4.1, Tablo 4.2) Google, Bing ve Yandex çevirileri de yapıldıktan sonra tüm makine çevirisi çıktıları A ve B çevirmenleri tarafından değerlendirilmek üzere çeviri ölçeğinde (Ek: 1 - Ek: 10) bir araya getirilmiştir. Dört farklı çeviri çıktısının hangi makine çevirisine ait olduğu, değerlendirici olan A ve B çevirmenlerine bildirilmemiştir. Çevirmenler tarafından Tablo 4.3 ile gösterilen derecelere göre değerlendirilen çıktılara karşılık gelen ağırlık puanı, cümlelerin değerlendirme puanı olarak atanmış ve bu ağırlık puanlarının aritmetik ortalaması alınarak her çeviri modeli için 0 ile 100 aralığında belirlenmiş bir çeviri skoru elde edilmiştir.

Öznel değerlendirme olarak da bilinen insan değerlendirmesi yönteminde özellikle şu üç noktaya dikkat edilmelidir: (i) HD çevirisinin KD cümlesine göre anlamsal ve sözdizimsel doğruluğu; (ii) cümlenin anlamsal olarak doğrudurken, sözdizimsel olarak yanlış olması; (iii) ve cümlenin hem anlamsal hem de sözdizimsel olarak yanlış olması (Bangalore, Bordel ve Riccardi, 2001). Değerlendirme sıralaması yaklaşımında akıcılık (*fluency*) ve uygunluk (*adequacy*) ölçümleri gerekli açıklamalarla birlikte isteğe göre beş, altı veya yedi katmanlı da olabilmektedir (Snover vd., 2009b) Cümlelerin akıcılık ve uygunluk değerleri, sözdizimsel ve

anlamsal doğruluğa bağlı olarak değişmektedir. Bahsedilen kriterlere göre tasarlanmış dört katmanlı çevirmen derecelendirme ölçütü örneği Tablo 4.3 ile aşağıda sunulmuştur.

**Tablo 4.3: Çevirmen derecelendirme ölçütü**

Derece	Açıklama	Ağırlık
<b>Tam</b>	Anlamsal ve sözdizimsel olarak doğru.	<b>100</b>
<b>Makul</b>	Anlamsal olarak doğru; sözdizimsel olarak yanlış.	<b>70</b>
<b>Eksik</b>	Anlamsal olarak yanlış; sözdizimsel olarak doğru.	<b>35</b>
<b>Yanlış</b>	Anlamsal ve sözdizimsel olarak yanlış.	<b>0</b>

Eş yazımlı kelimeler kural olarak buldukları isim veya fiil ifadelerinde kullanılma durumlarına göre farklı anlamlara sahip olabilir. Bu değişikliğin herhangi bir dış müdahale olmaksızın makine yoluyla tespit edilerek çevirisinin yapılabilmesi beklenir. Bu maksatla yukarıda (Tablo 4.3) belirtilen ölçütlere dayanan bir değerlendirme ölçeği (Ek: 1 - Ek: 10) oluşturulmuştur. Buna göre, eş yazımlı bir kelimenin kullanıldığı isim veya fiil ifadesinin kabul edilir düzeyde doğru çevrilmiş olması durumunda cümlenin çeviri skoru 100 (yüz); anlamsal olarak kabul edilebilir olduğu halde sözdizimsel olarak yanlış veya genel kabul gören kelime(ler) kullanılmamış ise 75 (yetmiş beş); anlamsal olarak yanlış ancak sözdizimsel olarak doğru sıralamayla kelime(ler) kullanılmış ise 35 (otuz beş); hem anlamsal hem de sözdizimsel olarak yanlış ise 0 (Sıfır) olarak belirlenmiştir.

#### **4.4. Çevirmen Değerlendirme Puanları**

Eş yazımlı bulduran cümlelerin Google, Bing, Yandex ve KMMÇ modeli makine çevirileri için çevirmen değerlendirme puanlaması yoluyla elde edilen çeviri skorları ve her bir modelin çeviri başarı değerlendirmeleri Tablo 4.4 ile aşağıda sunulmuştur.

**Tablo 4.4: Google, Bing, Yandex ve KMMÇ modeli çevirmen değerlendirme puanları ve her bir modelin çeviri skorlarını gösterir tablo**

(A) EŞ YAZIMLI KELİMELERİN İSİM İFADELERİNDE KULLANILDIĞI CÜMLE ÇEVİRİSİ PUANLAMALARI					(B) EŞ YAZIMLI KELİMELERİN FİİL İFADELERİNDE KULLANILDIĞI CÜMLE ÇEVİRİSİ PUANLAMALARI				
Madde	Google Skoru	Bing Skoru	Yandex Skoru	KMMÇ Skoru	Madde	Google Skoru	Bing Skoru	Yandex Skoru	KMMÇ Skoru
1.	100	100	100	100	41.	100	0	70	70
2.	100	100	100	100	42.	35	35	35	100
3.	100	100	100	100	43.	0	0	0	100
4.	100	100	100	100	44.	100	100	100	70
5.	100	100	100	100	45.	100	100	100	100
6.	100	100	100	70	46.	100	100	100	100
7.	100	100	100	100	47.	70	0	0	100
8.	100	100	0	100	48.	100	100	100	100
9.	100	100	0	100	49.	0	0	0	100
10.	100	100	100	100	50.	100	100	100	100
11.	100	100	100	100	51.	100	100	100	100
12.	100	100	100	100	52.	100	100	70	100
13.	100	100	100	100	53.	100	100	100	100
14.	100	100	100	100	54.	100	100	100	100
15.	100	100	100	100	55.	100	35	100	100
16.	70	100	100	100	56.	100	35	100	0
17.	100	100	100	100	57.	100	100	100	35
18.	100	100	100	100	58.	100	100	100	100
19.	100	100	100	100	59.	100	100	0	100
20.	100	100	100	100	60.	100	0	100	100
21.	100	100	100	0	61.	100	100	100	100
22.	100	100	100	100	62.	100	100	100	70
23.	100	100	100	100	63.	100	70	100	100
24.	100	70	100	0	64.	100	100	100	100
25.	100	100	100	100	65.	100	100	70	100
26.	100	100	100	100	66.	100	0	0	100
27.	100	100	100	100	67.	100	100	70	100
28.	100	100	100	100	68.	100	100	100	100
29.	100	100	100	35	69.	100	100	100	100
30.	0	100	0	100	70.	100	100	100	100
31.	100	100	100	100	71.	100	100	100	100
32.	0	0	0	100	72.	100	100	100	100
33.	100	100	100	100	73.	100	100	100	70
34.	100	100	100	100	74.	0	0	0	0
35.	70	70	100	100	75.	100	100	100	100
36.	100	100	100	100	76.	100	100	100	100
37.	100	100	100	100	77.	0	0	0	0
38.	100	100	100	100	78.	100	100	35	100
39.	100	100	100	100	79.	100	100	100	100
40.	100	100	100	100	80.	100	100	100	100
<b>AO:</b>	<b>93,5</b>	<b>96</b>	<b>90</b>	<b>92,63</b>	<b>AO:</b>	<b>87,63</b>	<b>74,38</b>	<b>76,25</b>	<b>87,88</b>

Çevirmen değerlendirme puanları tablosuna (Tablo 4.4) göre 80 maddeden 49'unda Google, Bing, Yandex ve KMMÇ çevirilerinin tam başarı gösterdiği, 29 maddede muhtelif hataların bulunduğu ve 2 maddede ise dört makine çevirisinin de başarısız olduğu gözlenmiştir.

Tablo 4.4 incelendiğinde eş yazımlı kelimelerin isim ifadelerinde kullanıldığı cümlelerin (A bölümü) çeviri başarı puanlamalarının fiil ifadelerinde kullanılan cümlelerin (B bölümü) çeviri başarı puanlamalarından daha yüksek olduğu fark edilmektedir. A bölümü çeviri skorlarının Google için 93.5, Bing için 96, Yandex için 90 ve KMMÇ için 92.63; B bölümü çeviri skorlarının ise Google için 87.63, Bing için 74.38, Yandex için 76.25 ve KMMÇ için 87.88 olduğu tespit edilmiştir.

Genelde çeviri başarıları veya başarısızlıklarının benzerlikler gösterdiğine işaret eden Tablo 4.4 verilerine göre A bölümündeki 6, 8, 9, 16, 21, 24, 29, 30, 32, 35'inci ve B bölümündeki 41, 42, 43, 44, 47, 49, 52, 55, 56, 57, 59, 60, 62, 63, 65, 66, 67, 73, 74, 77, 78'inci maddelerin çevirilerinde çeviri motorları arası farklılıklar olduğu ve bu farklılıkların ekseri küçük çeviri hataları olmakla birlikte aralarında dikkate değer derecede yanlış çevirilerin de olduğu gözlemlenmiştir. Bu nedenle söz konusu yanlış çeviri çıktılarının tablolar halinde ele alınarak çeviri hatalarının incelenmesi Tablo 4.4'ün daha anlamlı hale gelmesinde faydalı olacaktır.

#### **4.5. Eş Yazımlı Kelime Bulunduran Çeviri Hatalarının İncelenmesi**

Tablo 4.4'teki değerlendirme puanları düşük veya sıfır olan maddeler incelendiğinde çeviri hatalarının üç farklı durumda gerçekleştiği görülmektedir. Bu hatalar sırasıyla (i) eş yazımlı kelimenin yaygın kullanılan yanlış çevirisinin az kullanılan doğru çevirisine tercih edildiği; (ii) eş yazımlı kelime belirsizliğinin (*Homonymous ambiguity*) olduğu ve (iii) ek belirsizliğinin bulunduğu (*Attachment Ambiguity*) durumlarda ortaya çıkmaktadır. Bu üç farklı çeviri hatası durumu alt başlıklarda tablolar eşliğinde ele alınmıştır.

#### 4.5.1. Yaygın ifadenin az kullanılan ifadeye tercih edilmesi

Alsohybe'ye (2017) göre kelime veya ifadelerden biri eğitim verisinde daha çok kullanılmışsa çeviri motoru, yanlış çeviri de olsa, çok kullanılanı doğru çeviriye tercih edecektir. Google, Bing, Yandex ve KMMÇ çeviri çıktılarında (Ek: 11 - Ek: 14) bu duruma örnek gösterilebilecek cümleler için Tablo 4.5 incelenebilir.

**Tablo 4.5: Eş yazımlı kelimenin yaygın kullanılan anlamının seyrek kullanılan anlamına tercih edildiği çeviri örnekleri**

MADDE	İNGİLİZCE CÜMLE	GOOGLE	BING	YANDEX	KMMÇ
32.	The <i>does</i> are walking in the forest.	Ormanda yürüyorlar.	Ormandakiler yürüyor.	Does ormanda yürüyor.	<b>Ormanda geyikler yürümekte.</b>
43.	They are going to <i>beef</i> .	Sığır edecekler.	Sığır etiyle kavga ediyorlar.	Sığır eti yiyecekler.	<b>Sızlanacaklar.</b>
47.	The people <i>carp</i> about junk mail.	İnsanlar önemsiz posta hakkında konuşuyorlar.	İnsanlar önemsiz postalar hakkında sazanı.	İnsanlar önemsiz posta hakkında sazan.	<b>İnsanlar istenmeyen posta hakkında mızızlanıır.</b>
66.	The waves are going to <i>buffet</i> the shore.	<b>Dalgalar kıyıya vuracak.</b>	Dalgalar kıyıya açık büfe olacak.	Dalgalar kıyıya açık büfe olacak.	<b>Dalgalar sahili dövecek.</b>
74.	We are going to <i>park</i> in front of the TV.	Televizyonun önüne park edeceğiz.	Televizyonun önüne park edeceğiz.	Televizyonun önüne park edeceğiz.	Televizyon önünde park edeceğiz.

Tablo 4.5 için Google, Bing, Yandex ve KMMÇ çıktıları (Ek: 11 - Ek: 14) arasından söz konusu çeviri hatasına örnek gösterilebilecek maddelerden bazıları (32, 43, 47, 66, 74) seçilmiştir. Verilen İngilizce cümlelerde eş yazımlı kelimeler yatık; çevirisi doğru olan cümleler ise kalın harf karakterleri ile yazılmıştır.

Tablo 4.5 incelendiğinde madde 32’de “geyikler” anlamına gelen “*does*” kelimesinin “*does*” şeklinde; madde 43’te “sızlanmak” olarak çevrilmesi gereken “*beef*” kelimesinin “sığır”, “sığır eti” ve “sığır etiyle kavga” şeklinde; madde 47’de “mızızlanmak” veya “söylenmek” anlamına gelen “*carp*” kelimesinin “sazan” ve “sazanı” şeklinde; madde 66’da “vuracak” veya “dövecek” anlamında kullanılan “*buffer*” kelimesinin “büfe olacak” şeklinde ve son olarak madde 74’te “yerleşmek” veya “kurulmak” şeklinde çevrilmesi gereken “*park*” kelimesinin “park etmek” şeklinde çevrildiği görülmektedir.

Yukarıda (Tablo 4.5) geçen (yatık yazılı) eş yazımlı kelimelerin Tureng vb. tüm Türkçe-İngilizce sözlüklerde aratılması durumunda yaygın kullanımda olan çevirilerinin tabloda görülen hatalı kelime çevirilerine benzer olduğu görülecektir. (Çoğu çevrimiçi sözlükte Tureng’de olduğu gibi yaygın kullanılanlar kategorisi olmasa da kelimelerin çevirisi, yaygın kullanılanlardan seyrek kullanılanlara doğru sıralanır.) Yani, dört ayrı çeviri modeli de kelimeleri sözlüklerdeki popüler manalarına uygun şekilde çevirmekte ve bu durum Tablo 4.5’teki maddelerde geçen eş yazımlı kelimelerin yaygın kullanılan anlamının, seyrek kullanılan anlamına tercih edildiğine (eğitim verilerinde sık yan yana görülen kelimeler veya ifadelerin seçildiğine) işaret etmektedir. Benzer çeviri hatası Google, Bing, Yandex ve KMMÇ çıktılarındaki 9, 24, 41, 49, 51, 65, 77 ve 78’inci maddelerinde (Ek: 11 - Ek: 14) de tespit edilmiştir.

#### **4.5.2. Eş yazımlı belirsizliği (*Homonymous ambiguity*)**

Aslında bu belirsizlik durumu (*Homonymous Ambiguity*) eş yazımlı kelimelerin geneli için geçerli olmakla birlikte çeviri motorlarının eğitim verilerinin sınırlı olmasından ileri gelmektedir. Okpor’a (2014) göre bu tür çeviri hataları, doğru çevrilmesi beklenen kelime gruplarının eğitim verisinde büyük oranda yan yana görülmemiş olmasından kaynaklanır. Google, Bing, Yandex ve KMMÇ modeli çeviri çıktılarında (Ek: 11 - Ek: 14) bu duruma örnek gösterilebilecek cümleler Tablo 4.6 ile aşağıda sunulmuştur.

**Tablo 4.6: Eş yazımlı kelimenin ilişkilendirilecek anlamlarının fazla olmasından (Homonymous ambiguity) kaynaklanan yanlış çeviri örnekleri**

MADDE	İNGİLİZCE CÜMLE	GOOGLE	BING	YANDEX	KMMÇ
42.	They bear the pain.	Acı çekiyorlar.	Acı çekiyorlar.	Acıyı taşıyorlar.	<b>Acıya katlanırlar.</b>
59.	Page the man in the hospital.	<b>Hastanedeki doktora çağrı yapın.</b>	<b>Hastanedeki doktoru çağrıla.</b>	Hastanede doktor Page.	<b>Hastanedeki doktoru çağır.</b>
60.	They still pine for their lost wealth.	<b>Hala kayıp servetleri için özlemlerini sürdürüyorlar.</b>	Hala kayıp servetleri için tırnıyorlar.	<b>Hala kayıp servetleri için can atıyorlar.</b>	<b>Kayıp servetleri için hala özlem çekerler.</b>

Tablo 4.6 için dört makine çevirisi çıktıkları (Ek: 11 - Ek: 14) arasından 42, 59 ve 60'ncı maddeler seçilmiştir. Verilen İngilizce cümlelerde eş yazımlı kelimeler yatık; çevirisi doğru olan cümleler ise kalın harf karakterleri ile yazılmıştır.

Yukarıdaki Tablo 4.6 incelendiğinde madde 42'de "katlanmak" olarak çevrilmesi gereken "bear" kelimesinin "taşımak" veya "çekmek" şeklinde; madde 59'da "çağrı yapmak", "çağrılmak" veya "çağırılmak" anlamında kullanılan "page" kelimesinin "page" şeklinde; madde 60'da "özlemini çekmek" manasında kullanılan "pine" kelimesinin "tırnıyor" şeklinde çevrildiği görülmektedir. Ayrıca; Tablo 4.6'da verilmemiş olsa da benzer çeviri hatası Google, Bing, Yandex ve KMMÇ modeli çıktılarının 25, 29, 52 ve 63'üncü maddelerinde (Ek: 11 - Ek: 14) de tespit edilmiştir.

#### 4.5.3. Ek belirsizliği (*Attachment ambiguity*)

Çeviri çıktılarında belirsizlik (*Ambiguity*) kelime, ifade veya cümlede birden fazla farklı anlamların olması durumudur. İki temel belirsizlik türü vardır: Sözcüksel (tek bir kelimedede iki veya daha fazla anlam) ve sözdizimsel (tek bir cümle veya ifadede iki veya daha fazla anlam) belirsizlikler (Jha, 2017). Sözcüksel belirsizlik için Nordquist'in (2016) makalesinde şu örneğe yer verilmiştir: "We saw her duck."

cümlesi hem “Onun eğildiğini gördük.” hem de “Onun ördeğini gördük.” şeklinde çevrilebilir (James vd., 2007). Burada eklemek gerekir ki eş sesliler veya eş yazımlılar belirsizliği (*Homonymous Ambiguity*) olarak bilinen bu durum mevcut tez çalışmasının genel problem sorusunu ilgilendirmektedir.

Ancak, mezkur başlıkla alakalı hata yukarıda bahsedilen sözcüksel belirsizliğe değil; **sözdizimsel belirsizlik** türüne işaret etmektedir. Ek belirsizliği (*Attachment ambiguity*) olarak da bilinen bu tür anlam karmaşıklıkları, farklı görevlerde kullanılabilen ‘in’, ‘on’, ‘at’ gibi edatlardan (*Prepositions*) kaynaklanmaktadır. Örneğin; “*The girl found a book on Main Street.*” cümlesindeki ‘on’ edatı belirsizlik içerir. Kitabın Main Street’te bulunduğu veya kitabın konusunun Main Street üzerine yazılmış olup olmadığı konusunda kesin bir hükme varılamaz (Jha, 2017). Google, Bing, Yandex ve tez modeli makine çevirilerinde (Ek: 11 - Ek: 14) bu belirsizlik durumuna örnek gösterilebilecek cümleler Tablo 4.7 ile sunulmuştur.

**Tablo 4.7: Ek belirsizliklerine (Attachment ambiguity) dayanan yanlış çeviri örnekleri**

MADDE	İNGİLİZCE CÜMLE	GOOGLE	BING	YANDEX	KMMÇ
6.	The mailman kept the letter in a <i>box</i> in his room.	Postacı mektubu odasında bir kutuda sakladı.	Postacı mektubu odasındaki bir kutuda saklıyordu.	Postacı mektubu odasındaki bir kutuda sakladı.	Postacı odasındaki bir kutudaki mektubu tuttu.
44.	They <i>boil</i> the goat milk in a pan.	Keçi sütünü tavada kaynatırlar.	Keçi sütünü tavada kaynatıyorlar.	Keçi sütünü bir tavada kaynatırlar.	Bir tavadaki keçi sütünü kaynatırlar.
73.	They <i>park</i> their cars in the building.	Arabaları binaya park ediyorlar.	Arabaları binaya park ediyorlar.	Arabaları binaya park ediyorlar.	Binadaki arabaları park ederler.

Tablo 4.7 için Google, Bing, Yandex ve KMMÇ modeli çıktıları (Ek: 11 - Ek: 14) arasından söz konusu çeviri hatasına örnek gösterilebilecek maddelerden bazıları (6, 44, 73) seçilmiştir. Verilen İngilizce cümlelerde eş yazımlı kelimeler yatık; çevirisi doğru olan cümleler ise kalın harf karakterleri ile yazılmıştır.

Yukarıdaki (Tablo 4.7) üç farklı İngilizce çeviri girdisinde de kullanılmış olan 'in' edatı, ek belirsizliği (*Attachment ambiguity*) durumuna neden olmuş ve bu yüzden KMMÇ modelinin madde 6 çevirisinde "...kutuda mektubu..." şeklinde olması gereken ifade "...kutudaki mektubu..."; madde 44 çevirisinde "...bir tavada..." şeklinde olması gereken ifade "...bir tavadaki..." ve madde 73 çevirisinde "Arabaları binaya..." şeklinde olması gereken ifade "Binadaki arabaları..."; olarak Türkçeye çevrilmiştir. Google, Bing ve Yandex çevirilerinde bu hata türünün (80 madde içerisinde) görülmediği tespit edilmiştir.

Bir kelimenin anlamını kullanıldığı çerçevenin sınırları belirler. Henrik Nikula'nın (1986) ifade ettiği gibi: "*Bağlam, kelimelerin bir araya gelmesiyle oluşur. Bu nedenle metindeki bir kelimenin anlamına karar vermeden önce onu çevreleyen kelimelere bakılmalıdır. Metin dışı bilgiler bundan sonra gelir.*" Ek belirsizliği (*Attachment ambiguity*) durumlarında Google, Bing ve Yandex çevirilerinin KMMÇ modelinden daha başarılı çıktılar ortaya koyuyor olması kelimelerin yan yana görülme sıklığına dayanan algoritmalara sahip sistemler oldukları gerçeğiyle açıklanabilir.

## BEŞİNCİ BÖLÜM

### SONUÇLAR ve ÖNERİLER

Bu tez çalışmasında genel problem cümlesi olarak “Makine çevirilerinde sorun teşkil eden eş yazımlı kelime bulunduran KD ifadelerinin sözdizimsel ve anlambilimsel kurallara göre tasarlanmış bir çeviri modeli ile doğru şekilde hedef dile çevrilebilmesi mümkün müdür?” seçilmiştir. Bu amaçla tez çalışmasında önerilen kural tabanlı bir makine çevirisi modeli geliştirilmiş ve sunulmuştur. Ancak; önerilen KMMÇ modeli, bütünüyle çeviri problemlerine odaklanan ve kendisinden önceki makine çevirisi çalışmalarının kapatamadığı açıkları kapatma motivasyonu ile tasarlanmamıştır.

Çalışmada temel fikir olarak İMÇ’lerdeki (2.5.2) hiyerarşik ifade tabanlı istatistiksel makine çevirisine (Chiang, 2007) benzer bir model geliştirilmiş; ancak geliştirilen sisteminin eğitiminde Chiang’ın modelinin aksine paralel metinlerden yararlanılmayıp çevirmen tecrübelerine dayanan (kural tabanlı) bir yöntem tercih edilmiştir. Sunulan KMMÇ modelinde, başarı ölçümleri için, eş yazımlı kelimeler (*homonymous ambiguity*) bulunduran cümle çevirileri yapılmış; ardından ise elde edilen çıktıların Google, Bing ve Yandex çevirileri ile karşılaştırılması işlemi gerçekleştirilmiştir.

KMMÇ sisteminin karşılaştırıldığı diğer modellerden farklı olarak kural tabanlı bir yapıya sahip olması ve İngilizce-Türkçe gibi birbirinden uzak dil yapılarına sahip dillerin çevirisi için tasarlanmış olması nedeniyle kara kutu (*black box*) testi uygulanmıştır. White’a (2003) göre kara kutu testleri, araştırılan sistemin yapısını veya tasarım felsefesini dışarıda tuttuğu için farklı çeviri sistemlerinin karşılaştırılmasında tercih edilmelidir.

Karşılaştırılacak çeviri modelleri için en iyi değerlendirme yöntemi olarak kabul edilen çevirmen değerlendirmesinden yararlanılmıştır. Bunun için 80 maddeden oluşan bir çeviri değerlendirme ölçeği (Ek: 1 - Ek: 10) oluşturulmuş; bu ölçek yardımıyla A ve B çevirmenleri tarafından her bir makine çeviri çıktısı (Ek: 11 - Ek:

14) için çevirmen değerlendirme puanı ve her bir ayrı model için çeviri skoru (Tablo 4.4) elde edilmiştir.

Çalışmanın iki alt probleminden birincisi “İngilizcede isim ifadelerinde (*noun phrase*) geçen eş yazımlıların Türkçeye doğru şekilde çevrilebilmesi mümkün müdür?” sorusudur. Bu eş yazımlı kelimelerin isim ifadelerinde kullanıldığı cümle (Madde 1 – 40) çevirilerinin Google, Bing, Yandex, KMMÇ modeli çeviri skorları ve KMMÇ ile diğer modellerin çeviri skoru karşılaştırmaları için Tablo 5.1 aşağıda sunulmuştur.

**Tablo 5.1: Eş yazımlı kelimelerin isim ifadelerinde kullanıldığı cümle çevirilerinin Google, Bing, Yandex, KMMÇ modeli çeviri skorları ve KMMÇ ile diğer modellerin çeviri skoru karşılaştırması**

Madde 1 – 40 Çeviri Skorları		KMMÇ	KMMÇ'nin Diğer Modellere Göre Başarı Farkı Yüzdeleri
Google	93,5		Google ile %-0.93
Bing	96	92.63	Bing ile %-3.51
Yandex	90		Yandex ile %2.92

Tablo 5.1’de eş yazımlı kelimelerin isim ifadelerinde kullanıldığı cümle (Madde 1 – 40) çevirilerinin modellere göre çeviri skorları (Tablo 4.4) dikkate alındığında birinci sırada 96 ile Bing’in; ikinci sırada 93,5 ile Google’ın; üçüncü sırada 92,63 ile KMMÇ’nin ve dördüncü sırada ise 90 puan ile Yandex’in olduğu görülmektedir. Bu verilere göre tez çalışmasında önerilen KMMÇ modeli başarı farkı yüzdelerinin Google ile mukayese edildiğinde %-0,93; Bing ile mukayese edildiğinde %-3,51 ve Yandex ile mukayese edildiğinde ise %2,92 olduğu hesaplanmıştır.

Çalışmanın iki alt probleminden ikincisi “İngilizcede fiil grubu (*verb phrase*) olarak kullanılan ifadelerde geçen eş yazımlıların Türkçeye doğru şekilde çevrilebilmesi mümkün müdür?” sorusudur. Bu eş yazımlı kelimelerin fiil ifadelerinde kullanıldığı cümle (Madde 41 – 80) çevirilerinin Google, Bing, Yandex, KMMÇ modeli çeviri

skorları ve KMMÇ ile diğer modellerin çeviri skoru karşılaştırmaları için Tablo 5.2 incelenebilir.

**Tablo 5.2: Eş yazımlı kelimelerin fiil ifadelerinde kullanıldığı cümle çevirilerinin Google, Bing, Yandex, KMMÇ modeli çeviri skorları ve KMMÇ ile diğer modellerin çeviri skoru karşılaştırması**

Madde 41 – 80		KMMÇ	KMMÇ'nin Diğer Modellere Göre Başarı Farkı Yüzdeleri
Çeviri Skorları			
Google	87.63	87.88	Google ile %0.29
Bing	74.38		Bing ile %18.15
Yandex	76.25		Yandex ile %15.25

Tablo 5.2'de eş yazımlı kelimelerin fiil ifadelerinde kullanıldığı cümle (Madde 41 – 80) çevirilerinin modellere göre (Tablo 4.4) çeviri skorları dikkate alındığında birinci sırada 87,88 ile KMMÇ'nin; ikinci sırada 87,63 ile Google'ın; üçüncü sırada 76,25 ile Yandex'in ve son olarak dördüncü sırada 74,38 puan ile Bing'in olduğu görülmektedir. Buna göre, tez çalışmasında önerilen KMMÇ modeli başarı farkı yüzdelerinin Google ile mukayese edildiğinde %0,29; Bing ile mukayese edildiğinde %18,15 ve Yandex ile mukayese edildiğinde ise %15,25 olduğu tespit edilmiştir.

Çalışmada tüm eş yazımlı kelimelerin isim ve fiil ifadelerinde kullanıldığı cümle (Madde 1 – 80) çevirilerinin Google, Bing, Yandex, KMMÇ modeli çeviri skorları ve KMMÇ ile diğer modellerin çeviri skoru karşılaştırmalarının olduğu Tablo 5.3 aşağıda sunulmuştur.

**Tablo 5.3: Tüm eş yazımlı kelimelerin isim ve fiil ifadelerinde kullanıldığı cümle çevirilerinin Google, Bing, Yandex, KMMÇ modeli çeviri skorları ve KMMÇ ile diğer modellerin çeviri skoru karşılaştırması**

Madde 1 – 80 Çeviri Skorları		KMMÇ	KMMÇ'nin Diğer Modellere Göre Başarı Farkı Yüzdeleri
Google	90.57	90.26	Google ile %-0.34
Bing	85.19		Bing ile %5.95
Yandex	83.13		Yandex ile %8.58

Tablo 5.3'te tüm eş yazımlı kelimelerin hem isim hem fiil ifadelerinde kullanıldığı cümle (Madde 1 – 80) çevirilerinin modellere göre çeviri skorları (Tablo 4.4) dikkate alındığında birinci sırada 90,57 puan ile Google'ın; ikinci sırada 90,26 puan ile KMMÇ'nin; üçüncü sırada 85,19 puan ile Bing'in ve son olarak dördüncü sırada 83,13 puan ile Yandex'in olduğu görülmektedir. Bu veriler ışığında tez çalışmasında önerilen KMMÇ modeli başarı farkı yüzdelerinin Google ile mukayese edildiğinde %-0,34; Bing ile mukayese edildiğinde %5,95 ve Yandex ile mukayese edildiğinde ise %8,58 olduğu tespit edilmiştir.

Bu tez çalışmasında Google, Bing, Yandex ve KMMÇ modeli makine çevirilerinin başarı puanları karşılaştırmalarında anlamlı bir fark görülmemiş; ancak elde edilen tüm makine çevirisi hatalarının üç farklı noktada kümelendiği tespit edilmiştir. Uygulama bölümünde (4.5) bu üç farklı hata durumlarına tablolar eşliğinde değinilmiş olsa da kısaca şu şekilde özetlenebilir:

- i. **Eş yazımlı kelimenin yaygın kullanılan anlamının seyrek kullanılan anlamına tercih edilmesi:** Çeviri motorlarının eğitim verilerinde kullanılan çeviri çiftlerinde sık yan yana görülen kelimeler veya ifadelerin çeviri adayı olarak seçilmesinden kaynaklanan bir hata durumudur. Bu hata durumunu engellemek için eğitim verilerinin artırılması yoluna gidilse de kelime anlamlarının yaygın veya seyrek kullanım oranlarında çeviri çıktıklarına etki edecek düzeyde bir değişiklik oluşmaz. Yani "does"

eş yazımlısının metin içinde "geyikler" anlamında görülme sıklığı ile "yapar" anlamında görülme sıklığı, sistemin eğitilmesi için gerekli olan çeviri çiftlerinin artırılması durumunda da benzer oranda artış gösterecektir.

- ii. **Eş yazımlı belirsizliği (*Homonymous Ambiguity*):** Eş yazımlı kelime bulunduran tüm çeviri durumlarında karşılaşılabılır. Bu hata türü çeviri motorlarının eğitim verilerinin yetersiz olmasından kaynaklanmaktadır.
- iii. **Ek belirsizliği (*Attachment ambiguity*):** Farklı görevlerde kullanılabilen 'in', 'on', 'at' gibi edatlardan (*Prepositions*) kaynaklanmaktadır. Bu hata türü "The girl found a book on Main Street." cümlesinde kızın kitabı Main Street'te bulunduğu veya kitabın Main Street üzerine yazılmış olduğu gibi iki farklı anlamının olmasından kaynaklanan belirsizlik durumudur. Ek belirsizliği hata türünün önerilen çeviri modelinin çeviri hatalarının 1/3'üne (Madde 6, 21, 44 ve 73) tekabül ettiği; Google, Bing ve Yandex modellerinde ise bu hata türüne rastlanmadığı tespit edilmiştir.

KMMÇ modelinin çeviri skorunun diğer makine çevirileri arasında 90,26 ile ikinci sırada olması ve Google ile kıyasla %-0,34; Bing ile kıyasla %5,95 ve Yandex ile kıyasla %8,58 seviyelerinde başarı kaydetmiş olması dikkate değerdir. Ancak modelin eş yazımlı hatalarında minimum hata ile en iyi performansı gösterdiğini ve sıralamadaki ikinciliğinin, eş yazımlı hata türünden farklı bir hata türü olan "ek belirsizliği" (*Attachment ambiguity*) durumlarından kaynaklandığını belirtmekte yarar vardır.

McMillan'a (2000) göre makine çevirisi çalışmalarının değerlendirilmesinde en önemli unsur veri, veri analizi ve sonuçlarının inandırıcı ve güvenilir olmasıdır. Bu nedenle sunulmakta olan çeviri modeli çalışmasında hedef dil çıktılarında kontrol edilmesi gereken "yeterlilik" ve "akıcılık" noktalarının ölçülmesinde çevirmen değerlendirmesinden yararlanılmıştır. Çok daha az güvenilir olmasına rağmen makine çevirilerinde otomatik değerlendirme yöntemlerinin yaygın şekilde kullanılıyor olması hızlı ve düşük maliyet tercihlerinden kaynaklanmaktadır. Bu çalışmadaki gibi hız ve maliyet kaygısı bulunmayan ölçümlenmelere sahip

arařtırmalarda gereki sonular iin insan deęerlendirmesi ynteminin kullanılması yerinde olacaktır.

Otomatik eviri deęerlendirmelerinde kalite puanlandırmaları, makine ıktıları ile evirmenler tarafından evrilen bir veya birden fazla referans metinlerin karřılařtırılması yoluyla yapılır. Ancak, Trke gibi ekim-yapım ekleri olan morfolojik zenginlięe sahip eklemeli (*Agglutinative*) bir dilde, tek bir ıktı iin ok farklı kelime sıralamalarına sahip referans metinleri olabilir. Bu nedenle, otomatik eviri deęerlendirmelerinde Trke, Macarca, Japonca gibi eklemeli diller sz konusu olduęunda yapılacak puanlamaların ok daha zor olabileceęi ngrlebilir.

Doęal diller tam olarak matematiksel formllerle ifade edilemedikleri ve sosyal kabullere dayanan ifadeler barındırdıkları iin biimsel ve anlamsal tutarsızlıklar sergileyebilmektedir. zellikle eř yazımlı bulunduran cmlelerin bařka doęal dillere, doęru evirisi yalnız her iki dilin dilbilimsel ve kltrel derinliklerine hkim evirmenler tarafından gerekleřtirilebilir. Bugne kadar teknik imknların elverdięi lde yapılan tm makine evirisi alıřmalarının ancak insan evirilerine yakın ıktılar elde edebildięi grlmřtr. Bu nedenle makine evirilerinde hızlı ve kaliteli metin evirileri iin n dzenleme (*pre-editing*) ve son dzenleme (*post editing*) yardımıyla eksikliklerin kapatılması, halihazırda en doęru zm yolu olarak grnmektedir.

## KAYNAKÇA

- Adalı, E. (2012). Doğal Dil İşleme, *Türkiye Bilişim Vakfı Bilgisayar Bilimleri ve Mühendisliği Dergisi*, 5 (2)  
[https://scholar.google.com.tr/citations?view\\_op=view\\_citation&hl=tr&user=GX9eTvAAAAAJ&citation\\_for\\_view=GX9eTvAAAAAJ:mB3voiENLucC](https://scholar.google.com.tr/citations?view_op=view_citation&hl=tr&user=GX9eTvAAAAAJ&citation_for_view=GX9eTvAAAAAJ:mB3voiENLucC)  
[Erişim: 05.11.2020]
- Ahmed A. ve Hanneman G. (2005). Syntax-Based Statistical Machine Translation: A review, *Association for Computational Linguistics*, Carnegie Mellon University,  
[http://www.2.cs.cmu.edu/afs/cs.cmu.edu/project/cmt\\_55/lti/Courses/734/Spring-08/Amr+Greg-survey-SSMT.pdf](http://www.2.cs.cmu.edu/afs/cs.cmu.edu/project/cmt_55/lti/Courses/734/Spring-08/Amr+Greg-survey-SSMT.pdf) [Erişim: 18.10.2020]
- Aho, A. V. ve Ullman, J. D. (1969). Syntax directed translations and the pushdown assembler. *Journal of Computer and System Sciences*, 3:37–56.
- Alammar, J. (2018). *Visualizing machine learning one concept at a time*, <http://jalammar.github.io/illustrated-transformer/> [Erişim: 01.11.2020]
- Alsohybe N. T., Dahan N. A. ve Ba-Alwi F. M. (2017). Machine-Translation History and Evolution: Survey for Arabic-English Translations, *Current Journal of Applied Science and Technology* 23(4): 1-19, 2017; Article no.CJAST.36124.
- Antony P. J. (2013). Machine Translation Approaches and Survey for Indian Languages, *Computational Linguistics and Chinese Language Processing*, Vol. 18, No. 1, March
- Aslan, E. (2018). Otomatik Çeviri Araçlarının Yabancı Dil Öğretiminde Kullanımı: Google Çeviri Örneği, *SEFAD*, 2018 (39): 87-104 - e-ISSN: 2458-908X
- Aslan, E. (2019). *Makine Çevirisi*, Hiperlink Yayıncılık, İstanbul.
- Avramidis, E. ve Koehn, P. (2008). Enriching Morphologically Poor Languages for Statistical Machine Translation., *ACL*, s.763–770.

- Ba, J. L., Kiros, J. R. ve Hinton, G. E. (2016). Layer normalization, *arXiv preprint arXiv:1607.06450*, 2016. <https://arxiv.org/abs/1607.06450> [Eriřim: 30.04.2021]
- Bahdanau, D., Cho, K. ve Bengio, Y. (2014). Neural Machine Translation by Jointly Learning to Align and Translate. *arXiv preprint arXiv:1409.0473*.
- Bahdanau, D., Cho, K. and Bengio, Y. (2015). Neural Machine Translation by Jointly Learning to Align and Translate, *Proceedings of the Third International Conference on Learning Representations*. <https://arxiv.org/abs/1409.0473v6> [Eriřim: 21.03.2021]
- Balkul, H. İ. (2015). *Türkiye’de Akademik Çeviri Eğitiminde Çeviri Teknolojilerinin Yerinin Sorgulanması: Müfredat Analizi Ve Öğretim Elemanlarının Konuya İliřkin Görüşleri Üzerinden Bir İnceleme* (Doktora Tezi), Sakarya Üniversitesi, Sosyal Bilimler Enstitüsü, Sakarya.
- Ballabh, A. ve Jaiswal, U. C. (2015). A Study of Machine Translation Methods and their Challenges, *Int. J. Adv. Res. Sci. Eng.* 4(1): 423-429.
- Banerjee, S. ve Lavie, A. (2005). METEOR: An Automatic Metric for MT Evaluation with Improved Correlation with Human Judgments, *in ACL Workshop on Intrinsic and Extrinsic Evaluation Measures for Machine Translation and/or Summarization Ann Arbor, MI, USA*.
- Bangalore B., Bordel G. ve Riccardi G. (2001) Computing consensus translation from multiple machine translation systems *IEEE Workshop on Automatic Speech Recognition and Understanding, 2001. ASRU '01*, Madonna di Campiglio, Italy  
[https://scholar.google.com.tr/citations?view\\_op=view\\_citation&hl=th&user=k9pXsKoAAAAJ&citation\\_for\\_view=k9pXsKoAAAAJ:9yKSN-GCB0IC](https://scholar.google.com.tr/citations?view_op=view_citation&hl=th&user=k9pXsKoAAAAJ&citation_for_view=k9pXsKoAAAAJ:9yKSN-GCB0IC)  
[Eriřim: 20.04.2021]
- Bar-Hillel, Y. (1962). The Future of Machine Translation. *Times Literary Supplement, Times Newspapers*, April 20th. London.

*bing.com/translator* [Erişim: 01.05.2021]

Brown, P., Cocke, J., Della Pietra, S., Della Pietra, V., Jelinek, F., Mercer, F. ve Roossin, P. (1988). A statistical approach to French/English translation, *Proceedings of the Second International Conference on Theoretical and Methodological Issues in Machine Translation of Natural Languages*.

Brown, P., Cocke, J., Pietra, S. D., Pietra, V. D., Jelinek, F., Lafferty, J., Mercer, R., ve Roossin, P. (1990). A statistical approach to machine translation. *Computational Linguistics*, 16:79–85.

Brown, P., Pietra, V. D., Pietra, S. D. ve Mercer, R. (1993). The mathematics of statistical machine translation: parameter estimation. *Computational Linguistics*, 19: 263–311.

Castaño M. A, Casacuberta, F. ve Vidal, E. (1997). Machine translation using neural networks and finite-state models, *Published in TMI*, <https://pdfs.semanticscholar.org/ad2c/fa96bf9149c8d2d5be606edf4203933e6194.pdf>. [Erişim: 22.03.2021]

Callı, Ahmet (2016). *Makine Çevirileri*. <http://www.estr.com/index.php/tr/makine-cevirileri/> [Erişim Tarihi: 21.09.2020].

Callison-Burch, C. ve Osborne, M. (2006). Re-evaluating the role of BLEU in machine translation research, *11st Conference of the European Chapter of the Association for Computational Linguistics, Proceedings of the Conference, April 3-7, 2006, Trento, Italy*.

*ceviri.yandex.com.tr* [Erişim: 01.05.2021]

Cheng, J. Dong, L. ve Lapata, M. (2016). Long short-term memory-networks for machine reading, *arXiv preprint arXiv:1601.06733*

Chiang, D. (2005). A hierarchical phrase-based model for statistical machine translation. *Proceedings of the 43rd Annual Meeting of the ACL* (s. 263–270), Ann Arbor, MI

----- (2006). An Introduction to Synchronous Grammars, <https://www.semanticscholar.org/paper/An-Introduction-to-Synchronous-Grammars-Chiang/61863c024c81eb72d72068118765ba7dccc8c3e0> [Eriřim: 01.01.2021]

----- (2007). Hierarchical Phrase-Based Translation. *Computational Linguistics*, 33(2): 201-228.

Cho, K., van Merriënboer, B., Bahdanau, D., and Bengio, Y. (2014a). On the properties of neural machine translation: Encoder–decoder approaches, *Proceedings of the Eighth Workshop on Syntax, Semantics and Structure in Statistical Translation*, sayfa: 103–111, Doha, Qatar. Association for Computational Linguistics.

Cho, K. van Merriënboer, B., Gülçehre Ç., Bahdanau, D., Bougares, F. Schwenk, H. ve Bengio, Y. (2014b). Learning phrase representations using rnn encoder–decoder for statistical machine translation, *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, sayfa: 1724–1734, Doha, Qatar, October 2014. Association for Computational Linguistics. doi: 10.3115/v1/D14-1179. URL <https://www.aclweb.org/anthology/D14-1179>.

Chomsky, N. (1986). Knowledge of Language: Its Nature, Origin, and Use, *Convergence* (Praeger Publishers), P106.C518 1986 410 85-12234, ISBN 0-275-90025-8, New York.

*collinsdictionary.com* [Eriřim: 30.04.2021]

Çiçekli, Ğ. ve Güvenir, H. A., (2001). Learning Translation Templates from Bilingual Translation Examples, *Applied Intelligence*, 15, 1: 57-76.

*dictionary.cambridge.org* [Eriřim: 29.04.2021]

- Doddington, G., (2002). Automatic Evaluation of MT Quality Using n-gram Co-occurrence Statistics, *Proceedings of Human Language Technology Conference*, San Diego, CA, pp. 128–132.
- Drury, D. A. (1969). "Homographs and Pseudo-Homographs," *Word Ways*: Vol. 2:Iss. 3 , Article 10. <https://digitalcommons.butler.edu/wordways/vol2/iss3/10>
- Durgar El-Kahlout, İ. ve Oflazer K. (2006). "Türkçe-İngilizce için istatistiksel bilgisayarlı çeviri sistemi." *Fifteenth Turkish Symposium on Artificial Intelligence and Neural Networks*, Akyaka, Muğla, Türkiye.
- Durgar El-Kahlout, İ., (2009). *A Prototype English-Turkish Statistical Machine Translation System* (Doktora Tezi), Sabancı Üniversitesi, İstanbul.
- Elloumi, H. Blanchon, Serasset G. ve Besacier L. (2015). Meteor for multiple target languages using dbnary, *MT Summit 2015*.
- Galley, M., Graehl, J., Knight, K., Marcu, D., DeNeefe, S., Wang, W. ve Thayer, I. (2006). Scalable inference and training of context-rich syntactic translation models, *Proceedings of the 21st International Conference on Computational Linguistics and the 44th annual meeting of the Association for Computational Linguistics* (s. 961-968). Association for Computational Linguistics.
- Goodman, J. T. (2001). A Bit of Progress in Language Modeling Extended Version, *Machine Learning and Applied Statistics Group*, Microsoft Research, One Microsoft Way, Redmond, WA 98052, <http://www.research.microsoft.com>
- Gough N. ve Way A. (2004). Robust large-scale EBMT with marker-based segmentation. *Proceedings of the Tenth International Conference on Theoretical and Methodological Issues in Machine Translation*, 95–104. Baltimore, MD.
- Groves D. (2007). Hybrid Data-Driven Models of Machine Translation, *Article in Machine Translation* (Ph.D.), Dublin City University, School of Computing, Dublin, Ireland.

- Hanneman, G., Huber, E., Agarwal, A., Ambati, V. Parlikar, A. Peterson, E. ve Lavie, A. (2008). A Statistical Transfer Systems for French–English and German–English Machine Translation, *Proceedings of the Third Workshop on Statistical Machine Translation*, Columbus, Ohio, pp. 163- 166.
- He, K., Zhang, X., Ren, S., ve Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 770–778, 2016.
- Hochreiter, S. and Schmidhuber, J. (1997). Long Short-Term Memory, *Neural Computation*. Nov. 1997, 9.8, pp. 1735–1780.
- Huang, L., Knight, K., Joshi, A. (2006) Statistical syntax-directed translation with extended domain of locality. In *Proceedings of AMTA* (pp. 66-73).
- Hutchins, J. (2005). Example-based machine translation: a review and commentary, *Mach. Transl.*, 19, 3–4: 197–211.
- (2015). Machine Translation: History of Research and Applications, *The Routledge Encyclopedia of Translation Technology* (s. 120-136). Londra ve New York: Routledge Taylor & Francis Group
- James, R. H., Brendan H., and Michael B. S. (2007). *Semantics: A Coursebook*, 2nd ed. Cambridge University Press
- Jha S. K. (2017). Exploring Ambiguous Structures In English, *Global Journal of Interdisciplinary Social Sciences*, *G.J.I.S.S.*, 6(3):1-4, ISSN: 2319-8834
- Jurafsky D. ve Martin J.H., (2000). An Introduction to Natural Language Processing, *Computational Linguistics and Speech Recognition* : 577-583.
- , D. ve Martin, J. H. (2009). *Speech and Language Processing: An Introduction to Natural Language Processing, and Computational Linguistics*, 2nd edition, Prentice-Hall.

- Kaji, H., Kida Y. ve Morimoto Y. (1992). Learning translation templates from bilingual text, *Coling-92: Proceedings of the Fifteenth [sic] International Conference on Computational Linguistics*, Nantes, France; pp. 672-678.
- Kalchbrenner N. ve Blunsom P. (2013). Recurrent Continuous Translation Models, *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, pages 1700–1709, Seattle, Washington, USA. Association for Computational Linguistics. URL <https://www.aclweb.org/anthology/D13-1176>.
- Kituku, B., Muchemi, L. ve Nganga W. (2016). Framework for Many to One Machine Translation, *International Journal of Advanced Research in Computer Science and Software Engineering*, Volume 6, Issue 5, May.
- Knight, K. ve Koehn, P. (2003). What's New in Statistical Machine Translation <http://www.snlp.de/prescher/teaching/2007/MachineTranslation/> [Erişim: 12.10.2020]
- Koehn, P., Och, F. J. ve Marcu, D. (2003). Statistical phrase-based translation, *Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology-Volume 1* (pp. 48-54). Association for Computational Linguistics.
- Koehn, P., Federico, M., Shen, W., Bertoldi, N., Bojar, O., Callison-Burch, C., Cowan, B., Dyer, C., Hoang, H., Zens, R. ve diğerleri (2006). Open Source Toolkit For Statistical Machine Translation: Factored Translation Models and Confusion Network Decoding, *Final Report of the 2006 JHU Summer Workshop*.
- Koehn, P. (2007). Statistical Machine Translation, Part 1, <https://www.google.com/search?client=firefox-b-d&q=www.mt-archive.info+%E2%80%BA+MTS-2007-Koehn-3> [Erişim: 10.10.2020]

- Koehn, P. ve Hoang, H. (2007). Factored Translation Models, <https://www.researchgate.net/publication/221013253> *Conference Paper January 2007* [Eriřim: 21.10.2020]
- Koehn, P., Hoang, H., Birch, A., Callison-Burch, C., Federico, M., Bertoldi, N., Cowan, B., Shen, W., Moran, C., Zens, R. ve Dyer, C. (2007). Moses: Open source toolkit for statistical machine translation, *Proceedings of the 45th annual meeting of the ACL on interactive poster and demonstration sessions* (pp. 177-180). Association for Computational Linguistics.
- Koehn, P. ve Knowles, R. (2017). Six Challenges for Neural Machine Translation, *Proceedings of the First Workshop on Neural Machine Translation*, s: 28-39, Vancouver, Canada.
- Korkmaz, İ. (2019). Makine Çevirisinin Kısa Tarihi, *International Journal of Social Humanities Sciences Research (JSHSR)*, January 2019 DOI: 10.26450/jshsr.1030
- Lewis, P. M., II ve R. E. Stearns (1968). Syntax-directed transduction, *Journal of the ACM*, 15:465–488.
- Makoto Nagao (1984). A Framework of A Mechanical Translation Between Japanese and English by Analogy Principle, *Artificial and Human Intelligence* (A. Elithorn and R. Banerji, eds.). Elsevier Science Publishers, Amsterdam.
- Marcu D. ve Wong, W. (2002). A Phrase-based, Joint Probability Model for Statistical Machine Translation, *Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing - Volume 10, EMNLP '02*, pages 133–139, Stroudsburg, PA, USA Association for Computational Linguistics.
- McMillan, J. H. (2000). *Educational Research: Fundamentals for the Consumer*, Longman (Üçüncü Baskı), New York.

*macmillandictionary.com* [Eriřim: 29.04.2021]

Mermer, C. (2019). *Unsupervised Learning of Word Alignments for Statistical Machine Translation* (Doktora Tezi), Elektrik ve Elektronik MÜhendisliđi Bölümü, Bođaziçi Üniversitesi.

*merriam-webster.com* [Eriřim: 30.04.2021]

Mihrabi, Hüseyin (2006). *İngilizce-Türkçe (Machine Translation- MT)*. <http://www.sametran.com/index.php?content=ingilizceturkce> (Eriřim Tarihi: 21.09.2020).

Muegge, U. (2006). An Excellent Application for Crummy Machine Translation: Automatic Translation of a Large Database, Elisabeth Gräfe, editor, *Proceedings 55 of the Annual Conference of the German Society of Technical Communicators*, s. 18–21, Stuttgart: tekom.

Musaođlu, M. (2003). *Türkçede Çeviri ve Aktarma*, Ankara Üniversitesi, TÖMER, Ankara, Sayı 24: 1-22

Nießen, S., Och, F., Leusch, G. ve Ney, H. (2000). An Evaluation Tool for Machine Translation: Fast Evaluation for MT Research, *Proceedings of the 2nd International Conference on Language Resources and Evaluation (LREC)*, Athens, Greece, pp. 39–45.

Nikula, H. (1986). *Dependensgrammatik*. Malmö: Liber. s. 41

Nordquist, R. (2016). Ambiguity (language). <https://www.thoughtco.com/ambiguity-language-1692388> [Eriřim: 17.05.2021]

Nyberg, E. H. ve Mitamura, T. (1992). The KANT System: Fast, Accurate, High-Quality Translation in Practical Domains, *Proceedings of the 14th conference on Computational linguistics - Volume 3* (s. 1069-1073). Association for Computational Linguistics.

- Oflazer, K. (1994). Two-level Description of Turkish Morphology, *Literary and Linguistic Computing*, 9: 137-148.
- (2016). Türkçe ve Doğal Dil İşleme, *Türkiye Bilişim Vakfı Bilgisayar Bilimleri ve Mühendisliği Dergisi* Cilt 5, Sayı 2, 24.06.2016  
<https://dergipark.org.tr/pub/tbbmd/issue/22245/238795>
- Okpor, M. D. (2014). Machine Translation Approaches: Issues and Challenges, *IJCSI International Journal of Computer Science Issues*, Vol. 11, Issue 5, No 2, September 2014 / ISSN (Print): 1694-0814
- Oladosu, J. (2016). Approaches to Machine Translation: A Review, *FUOYE Journal of Engineering and Technology*, Volume 1, Issue 1, September.
- Olah, C. (2015). Understanding LSTM Networks, *Colah's Blog*, <http://colah.github.io/posts/2015-08-Understanding-LSTMs/> [Erişim: 01.11.2020]
- Orhun, M. (2010). *Uygurcadan Türkçeye Bilgisayarlı Çeviri* (Doktora Tezi), İstanbul Teknik Üniversitesi, F.B.E., Bilgisayar Müh. Anabilim Dalı, İstanbul.
- oxfordlearnersdictionaries.com* [Erişim: 29.04.2021]
- Papineni, K., Roukos, S., Ward, T. ve Wei-Jing Z. (2002). Bleu: a Method for Automatic Evaluation of Machine Translation, *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL)*, Philadelphia, s. 311-318.
- Poibeau, Thierry (2017). *Machine Translation*, The MIT Press Cambridge, Massachusetts, London, England
- Pym A. (2013). Translation Skill-Sets in a Machine-Translation Age, *Meta : journal des traducteurs / Meta: Translators' Journal*, vol. 58, n° 3, 2013, p. 487-503, <https://www.erudit.org/en/journals/meta/2013-v58-n3-meta01406/1025047ar/>

- Raffel, C. ve Ellis, D.P.W. (2015). Feed-forward networks with attention can solve some long-term memory problems. *arXiv preprint arXiv:1512.08756*.
- Reddy, M. V. ve Hanumanthappa, M. (2012). Interlingual Machine Translation. Chief Sahil Seth, editor, *Proc. of the Intl. Conf. on Advances in Computer Science and Electronics Engineering*. Universal Association of Computer and Electronics Engineers.
- Richardson, S.D., Vanderwende L. ve Dolan W. (1993). Combining dictionary-based and examplebased methods for natural language analysis, *TMI-93: the Fifth International Conference on Theoretical and Methodological Issues in Machine Translation*, Kyoto, Japan, July 14-16, 1993. Proceedings; pp.69-79.
- Rikters, M. (2019). *Hybrid Machine Translation by Combining Output from Multiple Machine Translation Systems* (Doctoral Thesis), Faculty of Computing, University of Latvia, Riga, Latvia.
- Saini S. ve Sahula V. (2015). A Survey of Machine Translation Techniques and Systems for Indian Languages, *2015 IEEE International Conference on Computational Intelligence & Communication Technology*, Ghaziabad, India. <https://ieeexplore.ieee.org/document/7078789> [Eriřim: 01.12.2020]
- Sakar, E. (2021). Trkede Eř Biimlilik Meselesi, *Dil ve Edebiyat Arařtırmaları (DEA)*, Bahar, 2021; (23) 405-426 ISSN: 1308-5069 - E-ISSN: 2149-0651
- Samedova, E. (2008). Eř Adlılık, Eř Adlılıęa Benzer Dil Olayları ve Trkiye Trkesindeki Basit Fiillerin Eř Adlılıęı, *Hazar niversitesi*, <http://dspace.khazar.org/handle/20.500.12323/798> [Eriřim: 20.05.2021]
- Shannon, C. (1948). A mathematical Theory of Communication. *Bell System Technical Journal* 27.379–423 and 623–656.
- Shannon, C. E., (2001). A mathematical theory of communication, *ACM SIGMOBILE Mobile Computing and Communications Review* 5.1 : 3-55.

- Silverman, D. (2001). *Interpreting qualitative data: Methods for analyzing talk, text and interaction*. London: SAGE Publication.
- Skadins, R. (2011). *Combined Use of Rule-Based and Corpus-Based Methods in Machine Translation* (Doctoral Thesis), University Of Latvia, Riga
- Snover, M. Madnani, N. Dorr, B. ve Schwartz, R. (2009a). TER-Plus: paraphrase semantic and alignment enhancements to Translation Edit Rate, Article in *Machine Translation* [https://www.researchgate.net/publication/225236506\\_TER\\_Plus\\_paraphrase\\_semantic\\_and\\_alignment\\_enhancements\\_to\\_Translation\\_Edit\\_Rate](https://www.researchgate.net/publication/225236506_TER_Plus_paraphrase_semantic_and_alignment_enhancements_to_Translation_Edit_Rate)
- Snover, M. Madnani, N. Dorr, B. ve Schwartz, R. (2009b). TER-Plus: paraphrase semantic and alignment enhancements to Translation Edit Rate, *Proceedings of the Fourth Workshop on Statistical Machine Translation*, s. 259–268, Athens, Greece, 30 March – 31 March 2009. c 2009 Association for Computational Linguistics.
- Somers, H. (1999). Review Article: Example-based Machine Translation. *Machine Translation*, 14(2):113–157.
- Somers, H. (2003). Introduction. H. Somers, ve H. Somers (Dü.) içinde, *Computers and Translation* (s. 1-11). Amsterdam / Philadelphia: John Benjamins Publishnig Company.
- Sumita, E., Furuse O. ve Iida H. (1993). An example-based disambiguation of prepositional phrase attachment'. In *TMI-93: the Fifth International Conference on Theoretical and Methodological Issues in Machine Translation*, Kyoto, Japan, July 14-16, 1993. Proceedings; s.80-91.
- Sutskever, I., Vinyals, O. ve Le, Q. V. (2014). Sequence to Sequence Learning with Neural Networks. In Ghahramani, Z., Welling, M., Cortes, C., Lawrence, N. D., and Weinberger, K. Q., editors, *Advances in Neural Information Processing Systems 27*, s. 3104–3112. Curran Associates, Inc.

- Şahin, M. (2013). *Çeviri ve Teknoloji*. Mengi Basım, İzmir.
- (2015). Çevirmen Adaylarının Gözünden İngilizce-Türkçe Bilgisayar Çevirisi ve Bilgisayar Destekli Çeviri: Google Deneyi. *Çeviribilim ve Uygulamalı Dergisi*, (21): 43-60.
- Tantuğ, A. C., (2007). *Akraba ve bitişken diller arasında bilgisayarlı çeviri için karma bir model* (Doktora Tezi), İstanbul Teknik Üniversitesi, Fenbilimleri Enstitüsü, İstanbul.
- Tantug, A.C., Oflazer, K. ve El-Kahlout, I.D. (2008). BLEU+: a Tool for Fine-Grained BLEU Computation., *LREC*.
- translate.google.com* [Erişim: 01.05.2021]
- Tripathi, S. ve Sarkhel, J. K. (2010). Approaches to MT, *Annals of Library and Information Studies*, 57: 388-393.
- Turian, J. P., Shen, L. ve Melamed, I. D. (2003). Evaluation of Machine Translation and its Evaluation, *MT Summit IX*, New Orleans, LA.
- Vandeghinste, V., P.Dirix ve I. Schuurman: (2005). Example-Based Translation without Parallel Corpora: First Experiment on a Prototype, MT Summit X, Phuket, Thailand, September 16, *Proceedings of Second Workshop on Example-Based Machine Translation*, s. 135-142.
- Vardar, B. (2002). *Açıklamalı Dilbilim Terimleri Sözlüğü*. İstanbul: Multilingual.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł. ve Polosukhin, I. (2017). Attention Is All You Need. *CoRR*. vol. abs/1706.03762. 1706.03762/<http://arxiv.org/abs/1706.03762>
- Vauquois, B. (1968). A survey of Formal Grammars and Algorithms for Recognition and Transformation in Mechanical Translation., *Ifip congress* (2): 1114–1122.

- Watanabe, H. (1993). A method of Extracting Translation Patterns from Translation Examples, *In TMI- 93: the Fifth International Conference on Theoretical and Methodological Issues in Machine Translation*, Kyoto, Japan, July 14-16, 1993. Proceedings; pp.292-301.
- White, J. (2003) How to Evaluate Machine Translation. in H. Somers (Ed.), *Computers and Translation*. Amsterdam: John Benjamins B. V.
- Weaver, W. (1949). Recent Contributions to the Mathematical Theory of Communication. In Shannon, C. E. and Weaver, W., editors, *The Mathematical Theory of Communication*, s. 94–117. The University of Illinois Press, Urbana, IL.
- (1955). Translation. In Locke, W. N., & Boothe, A.D. (Eds.) *Machine translation of languages* (s. 15–23). Cambridge, MA: MIT Press. Reprinted from a memorandum written by Weaver in 1949.
- (2020) American mathematician, *Yaz. Britannica editörleri*, <https://www.britannica.com/biography/Warren-Weaver> [Erişim: 03.10.2020]
- Wilks, Y. (2009). *Machine Translation: Its Scope and Limits*. 10.1007/978-0-387-72774-5, ISBN: 978-0-387-72774-5, Springer, Boston, MA.
- Wu D. (1997). Stochastic Inversion Transduction Grammars and Bilingual Parsing of Parallel Corpora, *Computational Linguistics*, 23:377–404.
- Yamada, K., & Knight, K. (2001). A syntax-based statistical translation model, *Proceedings of the 39th Annual Meeting on Association for Computational Linguistics* (s. 523–530). Toulouse, France.
- Young, T., Hazarika, D., Poria, S. ve Cambria, E. (2018). Recent Trends in Deep Learning Based Natural Language Processing. *IEEE Computational Intelligence Magazine*: 55-75.
- yourdictionary.com* [Erişim: 30.04.2021]

Zhang, H., Gildea, D. (2008). Efficient Multi-Pass Decoding for Synchronous Context Free Grammars, *ACL* (s. 209-217).

Zhou, P. Shi, W. Tian, J. Qi, Z. Li, B. Hao, H. ve Xu, B.(2016). Attention-based Bidirectional Long Short-term Memory Networks for Relation classification, *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics* (Volume 2: Short Papers), volume 2, pages 207–212.



## EKLER

### Ek: 1 - İsim İfadeleri Çeviri Değerlendirme Ölçeği (Madde 1 – 8)

Aşağıdaki İngilizceden Türkçeye çevrilmiş cümleleri verilen ölçek doğrultusunda değerlendiriniz. Anlamsal ve sözdizimsel olarak doğru ise " <b>TAM</b> "; anlamsal ancak sözdizimsel olarak yanlış ise " <b>MAKUL</b> "; anlamsal değil ama sözdizimsel olarak doğru ise " <b>EKSİK</b> "; hem anlamsal hem de sözdizimsel olarak yanlış ise " <b>YANLIŞ</b> " seçeneğini işaretleyiniz.			
CÜMLE	İNGİLİZCE CÜMLE	DÖRT FARKLI MAKİNE ÇEVİRİSİ ÇIKTISI	TAM MAKUL EKSİK YANLIŞ
1.	She was holding the baby in her arms.	Bebeği kucağında tutuyordu. Bebeği kucağında tutuyordu. Bebeği kucağında tutuyordu. Kollarında bebeği tutmaktaydı.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
2.	We saw the little bear cub.	Küçük ayı yavrusunu gördük. Küçük ayı yavrusunu gördük. Küçük ayı yavrusunu gördük. Küçük ayı yavrusunu gördük.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
3.	My friend is a beef producer.	Arkadaşım bir sığır eti üreticisi. Arkadaşım bir sığır eti üreticisi. Arkadaşım bir sığır eti üreticisi. Arkadaşım bir sığır eti üreticisi.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
4.	The boil on his arm was painful.	Kolundaki çıban acı vericiydi. Kolundaki çıban acı vericiydi. Kolundaki çıban acı vericiydi. Kolundaki çıban ağrı vericiydi.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
5.	He is reading a book about American history.	Amerikan tarihi hakkında bir kitap okuyor. Amerikan tarihi hakkında bir kitap okuyor. Amerikan tarihi hakkında bir kitap okuyor. Amerikan tarihi hakkında bir kitap okumakta.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
6.	The mailman kept the letter in a box in his room.	Postacı mektubu odasında bir kutuda sakladı. Postacı mektubu odasındaki bir kutuda saklıyordu. Postacı mektubu odasındaki bir kutuda sakladı. Postacı odasındaki bir kutudaki mektubu tuttu.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
7.	Carp is a fish.	Sazan bir balıktır. Sazan bir balıktır. Sazan bir balıktır. Sazan bir balık.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
8.	A clip from that video.	O videodan bir klip. O videodan bir klip. Bu video klibi. O görüntüden bir klip.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>

**Ek: 2 - İsim İfadeleri Çeviri Değerlendirme Ölçeği (Madde 9 – 16)**

Aşağıdaki İngilizceden Türkçeye çevrilmiş cümleleri verilen ölçek doğrultusunda değerlendiriniz. Anlamsal ve sözdizimsel olarak doğru ise " <b>TAM</b> "; anlamsal ancak sözdizimsel olarak yanlış ise " <b>MAKUL</b> "; anlamsal değil ama sözdizimsel olarak doğru ise " <b>EKSİK</b> "; hem anlamsal hem de sözdizimsel olarak yanlış ise " <b>YANLIŞ</b> " seçeneğini işaretleyiniz.			
CÜMLE	İNGİLİZCE CÜMLE	DÖRT FARKLI MAKİNE ÇEVİRİSİ ÇIKTISI	TAM MAKUL EKSİK YANLIŞ
9.	That bird is a crane.	O kuş bir turna. O kuş bir turna. O kuş bir vinç. O kuş bir turna.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
10.	An express shipment.	Ekspres kargo. Ekspres bir sevkiyat. Bir ekspres gönderi. Bir hızlı mal sevkiyatı.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
11.	The last drop of my coffee.	Kahvemın son damlası. Kahvemın son damlası. Kahvemın son damlası. Kahvemın son damlası.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
12.	My friends at work love fast cars.	İş yerindeki arkadaşlarım hızlı arabaları sever. İş arkadaşlarım hızlı arabaları sever. İş yerindeki arkadaşlarım hızlı arabalara bayılır. İşteki arkadaşlarım hızlı arabaları sever.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
13.	A green flag indicates the start of the race.	Yeşil bayrak yarışın başladığını gösterir. Yeşil bayrak yarışın başlangıcını gösterir. Yeşil bayrak yarışın başlangıcını gösterir. Yeşil bir bayrak yarışın başlangıcını gösterir.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
14.	A kid in high school.	Lisede bir çocuk. Lisede bir çocuk. Lisede bir çocuk. Lisedeki bir çocuk.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
15.	The children love to play in the leaves.	Çocuklar yapraklarda oynamayı severler. Çocuklar yapraklarda oynamayı sever. Çocuklar yapraklarda oynamayı severler. Çocuklar yapraklarda oynamayı sever.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
16.	She was very mean to her sister.	Kız kardeşi için çok kaba davrandı. Kız kardeşine çok kötü davranıyordu. Kız kardeşine çok kötü davrandı. Kız kardeşine çok kabaydı.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>

### Ek: 3 - İsim İfadeleri Çeviri Değerlendirme Ölçeği (Madde 17 – 24)

Aşağıdaki İngilizceden Türkçeye çevrilmiş cümleleri verilen ölçek doğrultusunda değerlendiriniz. Anlamsal ve sözdizimsel olarak doğru ise " <b>TAM</b> "; anlamsal ancak sözdizimsel olarak yanlış ise " <b>MAKUL</b> "; anlamsal değil ama sözdizimsel olarak doğru ise " <b>EKSİK</b> "; hem anlamsal hem de sözdizimsel olarak yanlış ise " <b>YANLIŞ</b> " seçeneğini işaretleyiniz.			
CÜMLE	İNGİLİZCE CÜMLE	DÖRT FARKLI MAKİNE ÇEVİRİSİ ÇIKTISI	TAM MAKUL EKSİK YANLIŞ
17.	Every move was painful for the old man.	Yaşlı adam için her hareket acı vericiydi. Her hareketi yaşlı adam için acı vericiydi. Her hareket yaşlı adam için acı vericiydi. Yaşlı adam için her hareket acı vericiydi.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
18.	A collection of precious objects.	Değerli nesnelere oluşan bir koleksiyon. Değerli nesnelere oluşan bir koleksiyon. Değerli eşyalardan oluşan bir koleksiyon. Değerli eşyaların bir koleksiyonu.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
19.	Fill in the form on the page.	Sayfadaki formu doldurun. Sayfadaki formu doldurun. Sayfadaki formu doldurun. Sayfadaki formu doldur.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
20.	A pine forest in the city.	Şehirde bir çam ormanı. Şehirde bir çam ormanı. Şehirde bir çam ormanı. Şehirdeki bir çam ormanı.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
21.	At the press of a button.	Bir düğmeye basarak. Bir düğmeye basarak. Bir düğmeye basarak. Bir düğmenin basımında.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
22.	The burning desert sun.	Yanan çöl güneşi. Yanan çöl güneşi. Yanan çöl güneşi. Yanan çöl güneşi.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
23.	Build a ship for the navy.	Donanma için bir gemi inşa et. Donanma için bir gemi inşa et. Donanma için bir gemi inşa et. Donanma için bir gemi inşa et.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
24.	I hate wearing a pink tie.	Pembe kravat takmaktan nefret ediyorum. Pembe kravat takmak istemem. Pembe kravat takmaktan nefret ediyorum. Pembe bir bağ giymekten nefret ederim.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>

**Ek: 4 - İsim İfadeleri Çeviri Değerlendirme Ölçeği (Madde 25 – 32)**

Aşağıdaki İngilizceden Türkçeye çevrilmiş cümleleri verilen ölçek doğrultusunda değerlendiriniz. Anlamsal ve sözdizimsel olarak doğru ise " <b>TAM</b> "; anlamsal ancak sözdizimsel olarak yanlış ise " <b>MAKUL</b> "; anlamsal değil ama sözdizimsel olarak doğru ise " <b>EKSİK</b> "; hem anlamsal hem de sözdizimsel olarak yanlış ise " <b>YANLIŞ</b> " seçeneğini işaretleyiniz.			
CÜMLE	İNGİLİZCE CÜMLE	DÖRT FARKLI MAKİNE ÇEVİRİSİ ÇIKTISI	TAM MAKUL EKSİK YANLIŞ
25.	They don't know the tie between my grandmother and the town.	Büyükanemle kasaba arasındaki bağı bilmiyorlar. Büyükanemle kasaba arasındaki bağı bilmiyorlar. Büyükanemle kasaba arasındaki bağı bilmiyorlar. Ninem ve şehir arasındaki bağı bilmezler.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
26.	Dinner will be a cold buffet.	Akşam yemeği soğuk açık büfe olacaktır. Akşam yemeği soğuk bir büfe olacak. Akşam yemeği soğuk büfe olacak. Akşam yemeği bir soğuk büfe olacak.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
27.	My uncle and my father are very close.	Amcam ve babam çok yakınlar. Amcam ve babam çok yakınlar. Amcam ve babam çok yakınlar. Amcam ve babam çok yakın.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
28.	The antique console will be in the living room.	Antik konsol oturma odasında olacak. Antika konsol oturma odasında olacak. Antika konsol oturma odasında olacak. Antika konsol salonda olacak.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
29.	This is the new word record.	Bu yeni dünya rekoru. Bu yeni dünya rekoru. Bu yeni dünya rekoru. Bu yeni dünya kayıdı.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
30.	I swatted the fly with a newspaper.	Sineği bir gazete ile karıştırdım. Sineği bir gazeteyle ezdim. Sineği bir gazete ile çırptım. Bir gazete ile sineğe vurdum.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
31.	The leaves rustle in the wind.	Yapraklar rüzgarda hışırdar. Yapraklar rüzgarda hışırdar. Yapraklar rüzgarda hışırdıyor. Yapraklar rüzgada hışırdar.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
32.	The does are walking in the forest.	Ormanda yürüyorlar. Ormandakiler yürüyor. Does ormanda yürüyor. Ormanda geyikler yürümekte.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>

**Ek: 5 - İsim İfadeleri Çeviri Değerlendirme Ölçeği (Madde 33 – 40)**

Aşağıdaki İngilizceden Türkçeye çevrilmiş cümleleri verilen ölçek doğrultusunda değerlendiriniz. Anlamsal ve sözdizimsel olarak doğru ise " <b>TAM</b> "; anlamsal ancak sözdizimsel olarak yanlış ise " <b>MAKUL</b> "; anlamsal değil ama sözdizimsel olarak doğru ise " <b>EKSİK</b> "; hem anlamsal hem de sözdizimsel olarak yanlış ise " <b>YANLIŞ</b> " seçeneğini işaretleyiniz.			
CÜMLE	İNGİLİZCE CÜMLE	DÖRT FARKLI MAKİNE ÇEVİRİSİ ÇIKTISI	TAM MAKUL EKSİK YANLIŞ
33.	They sat on the grass in the park.	Parktaki çimlere oturdular. Parkta çimlere oturdular. Parkta çimlerin üzerine oturdular. Parktaki çimende oturdular.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
34.	The gas stove was in the old house in the village.	Gaz sobası köydeki eski evdeydi. Gaz sobası köydeki eski evdeydi. Gaz sobası köydeki eski evdeydi. Gaz sobası köydeki eski evdeydi.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
35.	They will be live on the city radio.	Şehir radyosundan canlı yayınlanacaklar. Şehir radyosundan canlı yayınlanacaklar. Şehir radyosunda canlı yayın yapacaklar. Şehir radyosunda yayında olacaklar.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
36.	That person is really not your type.	Bu kişi gerçekten senin tipin değil. O kişi gerçekten senin tipin değil. O kişi senin tipin değil. Gerçekten o kişi tipiniz değil.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
37.	My roommates ate two jars of jam in two days.	Oda arkadaşlarım iki günde iki kavanoz reçel yedi. Oda arkadaşlarım iki günde iki kavanoz reçel yedi. Oda arkadaşlarım iki günde iki kavanoz reçel yedi. Oda arkadaşlarım iki günde iki kavanoz reçel yedi.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
38.	Endurance was the most important attribute in a teacher.	Dayanıklılık, bir öğretmendeki en önemli özellikti. Dayanıklılık bir öğretmendeki en önemli özellikti. Dayanıklılık bir öğretmenin en önemli özelliği idi. Bir öğretmende dayanıklılık en mühim vasıftı.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
39.	The old man's old watch is fast.	Yaşlı adamın eski saati hızlı. Yaşlı adamın eski saati hızlı. Yaşlı adamın eski saati çok hızlı. Yaşlı adamın eski saati hızlı.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
40.	There is a huge wave of divorce in the developed countries.	Gelişmiş ülkelerde büyük bir boşanma dalgası var. Gelişmiş ülkelerde büyük bir boşanma dalgası yaşanıyor. Gelişmiş ülkelerde büyük bir boşanma dalgası var. Gelişmiş ülkelerde büyük bir boşanma dalgası var.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>

**Ek: 6 - Fiil İfadeleri Çeviri Değerlendirme Ölçeği ( Madde 41 – 48)**

Aşağıdaki İngilizceden Türkçeye çevrilmiş cümleleri verilen ölçek doğrultusunda değerlendiriniz. Anlamsal ve sözdizimsel olarak doğru ise " <b>TAM</b> "; anlamsal ancak sözdizimsel olarak yanlış ise " <b>MAKUL</b> "; anlamsal değil ama sözdizimsel olarak doğru ise " <b>EKSİK</b> "; hem anlamsal hem de sözdizimsel olarak yanlış ise " <b>YANLIŞ</b> " seçeneğini işaretleyiniz.			
CÜMLE	İNGİLİZCE CÜMLE	DÖRT FARKLI MAKİNE ÇEVİRİSİ ÇIKTISI	TAM MAKUL EKSİK YANLIŞ
41.	She arms herself with knowledge.	Kendini bilgiyle donatıyor. Kendini bilgiyle silahlandırıyor. Kendini bilgiyle silahlandırıyor. Bilgi ile kendini silahlandırır.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
42.	They bear the pain.	Acı çekiyorlar. Acı çekiyorlar. Acıyı taşıyorlar. Acıya katlanırlar.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
43.	They are going to beef.	Sığır edecekler. Sığır etiyle kavga ediyorlar. Sığır eti yiyecekler. Sızlanacaklar.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
44.	They boil the goat milk in a pan.	Keçi sütünü tavada kaynatırlar. Keçi sütünü tavada kaynatıyorlar. Keçi sütünü bir tavada kaynatırlar. Bir tavadaki keçi sütünü kaynatırlar.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
45.	We book flights to big cities.	Büyük şehirlere uçuş rezervasyonu yapıyoruz. Büyük şehirlere uçuş rezervasyonu yapıyoruz. Büyük şehirlere uçuş rezervasyonu yapıyoruz. Büyük şehirlere uçuşlar rezerve ederiz.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
46.	He is going to box in the Olympics in summer.	Yaz aylarında olimpiyatlarda boks yapacak. Yaz aylarında olimpiyatlarda boks yapacak. Yaz aylarında olimpiyatlarda boks yapacak. Yazın Olimpiyatlarda dövüşecek.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
47.	The people carp about junk mail.	İnsanlar önemsiz posta hakkında konuşuyorlar. İnsanlar önemsiz postalar hakkında sazanı. İnsanlar önemsiz posta hakkında sazan. İnsanlar istenmeyen posta hakkında mızımızlanır.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
48.	Her mom is going to clip her hair.	Annesi saçını kesecek. Annesi saçını kesecek. Annesi saçını kesecek. Annesi saçını kesecek.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>

**Ek: 7 - Fiil İfadeleri Çeviri Değerlendirme Ölçeği (Madde 49 – 56)**

Aşağıdaki İngilizceden Türkçeye çevrilmiş cümleleri verilen ölçek doğrultusunda değerlendiriniz. Anlamsal ve sözdizimsel olarak doğru ise " <b>TAM</b> "; anlamsal ancak sözdizimsel olarak yanlış ise " <b>MAKUL</b> "; anlamsal değil ama sözdizimsel olarak doğru ise " <b>EKSİK</b> "; hem anlamsal hem de sözdizimsel olarak yanlış ise " <b>YANLIŞ</b> " seçeneğini işaretleyiniz.			
CÜMLE	İNGİLİZCE CÜMLE	DÖRT FARKLI MAKİNE ÇEVİRİSİ ÇIKTISI	TAM MAKUL EKSİK YANLIŞ
49.	They crane forward.	İleriye doğru döner. Öne doğru vinçle ilerliyorlar. Vinç ileri. İleriye boynunu uzatırlar.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
50.	I'd like to express my gratitude to everyone.	Herkese şükranlarımı sunmak isterim. Herkese şükranlarımı sunuyorum. Herkese şükranlarımı sunmak istiyorum. Herkese minnettarlığımı ifade etmek isterim.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
51.	She is going to drop the book.	Kitabı bırakacak. Kitabı bırakacak. Kitabı bırakacak. Kitabı düşürecek.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
52.	Muslims fast during Ramadan.	Müslümanlar Ramazan ayında oruç tutar. Müslümanlar Ramazan ayında oruç tutarlar. Ramazan'da oruç Müslümanlar. Ramazan süresince müslümanlar oruçlu tutar.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
53.	Flag the potential problems in our proposal.	Önerimizdeki olası sorunları işaretleyin. Teklifimizdeki olası sorunları işaretle. Teklifimizdeki olası sorunları işaretleyin. Teklifimizdeki muhtemel problemleri işaretle.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
54.	Don't kid yourself.	Kendinizi kandırmayın. Kendini kandırma. Kendini kandırma. Kendini kandırma.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
55.	Their father leaves home for work.	Babaları iş için evden ayrılıyor. Babaları işe giderken evden ayrılır. Babaları iş için evden ayrılıyor. İş için babaları evden ayrılır.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
56.	I didn't mean any harm.	Zarar vermek istemedim. Kötü bir şey demek istemedim. Kötü bir niyetim yoktu. Hiç zararı istemedim.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>

**Ek: 8 - Fiil İfadeleri Çeviri Değerlendirme Ölçeği (Madde 57 – 64)**

Aşağıdaki İngilizceden Türkçeye çevrilmiş cümleleri verilen ölçek doğrultusunda değerlendiriniz. Anlamsal ve sözdizimsel olarak doğru ise " <b>TAM</b> "; anlamsal ancak sözdizimsel olarak yanlış ise " <b>MAKUL</b> "; anlamsal değil ama sözdizimsel olarak doğru ise " <b>EKSİK</b> "; hem anlamsal hem de sözdizimsel olarak yanlış ise " <b>YANLIŞ</b> " seçeneğini işaretleyiniz.			
CÜMLE	İNGİLİZCE CÜMLE	DÖRT FARKLI MAKİNE ÇEVİRİSİ ÇIKTISI	TAM MAKUL EKSİK YANLIŞ
57.	They move the flag up and down.	Bayrağı yukarı ve aşağı hareket ettirirler. Bayrağı yukarı ve aşağı hareket ettiriyorlar. Bayrağı yukarı ve aşağı hareket ettirirler. Yukarı aşağı bayrağı taşırlar.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
58.	People at my work object to the book.	İşimdeki insanlar kitaba itiraz ediyor. İşyerimdeki insanlar kitaba itiraz etti. İş yerimdeki insanlar kitaba itiraz ediyor. İşimdeki insanlar kitaba itiraz eder.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
59.	Page the man in the hospital.	Hastanedeki doktora çağrı yapın. Hastanedeki doktoru çağrıyla. Hastanede doktor Page. Hastanedeki doktoru çağır.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
60.	They still pine for their lost wealth.	Hala kayıp servetleri için özlemlerini sürdürüyorlar. Hala kayıp servetleri için tırnıyorlar. Hala kayıp servetleri için can atıyorlar. Kayıp servetleri için hala özlem çekerler.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
61.	Press the button for the third floor.	Üçüncü katın düğmesine basın. Üçüncü kat için düğmeye basın. Üçüncü kat için düğmeye basın. Üçüncü kat için düğmeye bas.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
62.	He deserted his wife and children for another woman.	Karısını ve çocuklarını başka bir kadın için terk etti. Karısını ve çocuklarını başka bir kadın için terk etti. Başka bir kadın için karısını ve çocuklarını terk etti. Başka bir kadını için hanım ve çocuklarını terk etti.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
63.	The company ships its goods to twenty countries.	Şirket mallarını yirmi ülkeye gönderiyor. Şirket mallarını yirmi ülkeye gemi iletir. Şirket mallarını yirmi ülkeye gönderiyor. Şirket yirmi ülkeye malını yollar.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
64.	The robbers were going to tie the men to a pipe with cable.	Soyguncular, adamları kabloyla bir boruya bağlayacaklardı. Soyguncular adamları kabloyla boruya bağlayacaklardı. Soyguncular adamları kabloyla bir boruya bağlayacaklardı. Hırsızlar kablo ile bir boruya adamları bağlayacaktı.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>

**Ek: 9 - Fiil İfadeleri Çeviri Değerlendirme Ölçeği (Madde 65 – 72)**

Aşağıdaki İngilizceden Türkçeye çevrilmiş cümleleri verilen ölçek doğrultusunda değerlendiriniz. Anlamsal ve sözdizimsel olarak doğru ise " <b>TAM</b> "; anlamsal ancak sözdizimsel olarak yanlış ise " <b>MAKUL</b> "; anlamsal değil ama sözdizimsel olarak doğru ise " <b>EKSİK</b> "; hem anlamsal hem de sözdizimsel olarak yanlış ise " <b>YANLIŞ</b> " seçeneğini işaretleyiniz.			
CÜMLE	İNGİLİZCE CÜMLE	DÖRT FARKLI MAKİNE ÇEVİRİSİ ÇIKTISI	TAM MAKUL EKSİK YANLIŞ
65.	The buildings house the young students.	Binalar genç öğrencileri barındırıyor. Binalar genç öğrencileri barındırıyor. Binalar genç öğrencilere ev sahipliği yapıyor. Binalar genç öğrencileri barındırır.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
66.	The waves are going to buffet the shore.	Dalgalar kıyıya vuracak. Dalgalar kıyıya açık büfe olacak. Dalgalar kıyıya açık büfe olacak. Dalgalar sahili dövecek.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
67.	We close the office at noon on Fridays.	Cuma günleri öğle saatlerinde ofisi kapatıyoruz. Cuma günleri öğlen ofisi kapatıyoruz. Cuma günleri öğlen ofisini kapatıyoruz. Cuma günleri öğle vakti ofisi kapatırız.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
68.	She is going to console herself with the cat.	Kendini kediyle teselli edecek. Kediyle kendini avutacak. O kedi ile kendini teselli edecek. Kedi ile kendini avutacak.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
69.	These cameras record 25 frames per second.	Bu kameralar saniyede 25 kare kaydeder. Bu kameralar saniyede 25 kare kaydeder. Bu kameralar saniyede 25 kare kaydeder. Bu kameralar saniyede 25 kareyi kayda alır.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
70.	He was the first person to fly.	Uçan ilk kişiydi. Uçan ilk kişiydi. İlk uçan kişi oydu. ilk ucan kişiydi.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
71.	The little girls were going to fly the kite.	Küçük kızlar uçurtmayı uçuracaklardı. Küçük kızlar uçurtmayı uçuracaktı. Küçük kızlar uçurtmayı uçuracaktı. Küçük kızlar uçurtmayı uçuracaktı.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
72.	My daughter does her homework every night.	Kızım her gece ödevini yapıyor. Kızım her gece ödevini yapıyor. Kızım her gece ödevini yapıyor. Her gece kızım ev ödevini yapar.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>

**Ek: 10 - Fiil İfadeleri Çeviri Değerlendirme Ölçeği (Madde 73 – 80)**

Aşağıdaki İngilizceden Türkçeye çevrilmiş cümleleri verilen ölçek doğrultusunda değerlendiriniz. Anlamsal ve sözdizimsel olarak doğru ise " <b>TAM</b> "; anlamsal ancak sözdizimsel olarak yanlış ise " <b>MAKUL</b> "; anlamsal değil ama sözdizimsel olarak doğru ise " <b>EKSİK</b> "; hem anlamsal hem de sözdizimsel olarak yanlış ise " <b>YANLIŞ</b> " seçeneğini işaretleyiniz.			
CÜMLE	İNGİLİZCE CÜMLE	DÖRT FARKLI MAKİNE ÇEVİRİSİ ÇIKTISI	TAM MAKUL EKSİK YANLIŞ
73.	They park their cars in the building.	Arabaları binaya park ediyorlar. Arabaları binaya park ediyorlar. Arabaları binaya park ediyorlar. Binadaki arabaları park ederler.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
74.	We are going to park in front of the TV.	Televizyonun önüne park edeceğiz. Televizyonun önüne park edeceğiz. Televizyonun önüne park edeceğiz. Televizyon önünde park edeceğiz.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
75.	He wants to live the last years of his life with his dogs.	Hayatının son yıllarını köpekleriyle yaşamak istiyor. Hayatının son yıllarını köpekleriyle geçirmek istiyor. Hayatının son yıllarını köpekleriyle yaşamak istiyor. Köpekleri ile hayatının son yıllarını yaşamak ister.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
76.	I type over 100 words per minute.	Dakikada 100 kelimedenden fazla yazıyorum. Dakikada 100'den fazla kelime yazıyorum. Dakikada 100'den fazla kelime yazıyorum. Dakikada 100 kelime üzerinde yazarım.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
77.	Their man's gun will jam.	Adamlarının silahı sıkışacak. Adamlarının silahı çalacak. Adamlarının silahı sıkışacak. Adamlarlarının silahı sıkışacak.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
78.	She will attribute her success to her late father's hard work.	Başarısını rahmetli babasının sıkı çalışmasına bağlayacak. Başarısını rahmetli babasının sıkı çalışmasına bağlayacak. Başarısını geç babasının sıkı çalışmasına bağlayacak. Ölen babasının çok çalışmasına başarısını dayandıracak.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
79.	Our kids only watch YouTube videos at weekends.	Çocuklarımız YouTube videolarını sadece hafta sonları izliyor. Çocuklarımız YouTube videolarını sadece hafta sonları izliyor. Çocuklarımız sadece hafta sonları YouTube videolarını izliyor. Hafta sonları çocuklarımız sadece YouTube videolarını izler.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
80.	The people on the bus wave to the villagers.	Otobüsteki insanlar köylülere el sallıyor. Otobüstekiler köylülere el sallıyor. Otobüsteki insanlar köylülere el sallıyor. Otobüsteki insanlar köylülere el sallar.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>

**Ek: 11 - Google, Bing, Yandex ve KMMÇ çıktıları tablosu (Madde 1 – 21)**

Madde	İNGİLİZCE CÜMLE	GOOGLE ÇIKTISI	BING ÇIKTISI	YANDEX ÇIKTISI	KMMÇ ÇIKTISI
1	She was holding the baby in her <b>arms</b> .	Bebegi kucağında tutuyordu.	Bebegi kucağında tutuyordu.	Bebegi kucağında tutuyordu.	Kollarında bebeği tutmaktaydı.
2	We saw the little <b>bear cub</b> .	Küçük ayı yavrusunu gördük.	Küçük ayı yavrusunu gördük.	Küçük ayı yavrusunu gördük.	Küçük ayı yavrusunu gördük.
3	My friend is a <b>beef</b> producer.	Arkadaşım bir sığır eti üreticisi.	Arkadaşım bir sığır eti üreticisi.	Arkadaşım bir sığır eti üreticisi.	Arkadaşım bir sığır eti üreticisi.
4	The <b>boil</b> on his arm was painful.	Kolundaki çıban acı vericiydi.	Kolundaki çıban acı vericiydi.	Kolundaki çıban acı vericiydi.	Kolundaki çıban ağrı vericiydi.
5	He is reading a <b>book</b> about American history.	Amerikan tarihi hakkında bir kitap okuyor.	Amerikan tarihi hakkında bir kitap okuyor.	Amerikan tarihi hakkında bir kitap okuyor.	Amerikan tarihi hakkında bir kitap okumakta.
6	The mailman kept the letter in a <b>box</b> in his room.	Postacı mektubu odasında bir kutuda sakladı.	Postacı mektubu odasındaki bir kutuda saklıyordu.	Postacı mektubu odasındaki bir kutuda sakladı.	Postacı odasındaki bir kutudaki mektubu tuttu.
7	<b>Carp</b> is a fish.	Sazan bir balıktır.	Sazan bir balıktır.	Sazan bir balıktır.	Sazan bir balık.
8	A <b>clip</b> from that video.	O videodan bir klip.	O videodan bir klip.	Bu video klibi.	O görüntüden bir klip.
9	That bird is a <b>crane</b> .	O kuş bir turna.	O kuş bir turna.	O kuş bir vinç.	O kuş bir turna.
10	an <b>express</b> shipment.	Ekspres kargo.	Ekspres bir sevkiyat.	Bir ekspres gönderi.	Bir hızlı mal sevkiyatı.
11	The last <b>drop</b> of my coffee.	Kahvemın son damlası.	Kahvemın son damlası.	Kahvemın son damlası.	Kahvemın son damlası.
12	My friends at work love <b>fast</b> cars.	İş yerindeki arkadaşlarım hızlı arabaları sever.	İş arkadaşlarım hızlı arabaları sever.	İş yerindeki arkadaşlarım hızlı arabalara bayılır.	İşteki arkadaşlarım hızlı arabaları sever.
13	A green <b>flag</b> indicates the start of the race.	Yeşil bayrak yarışın başladığını gösterir.	Yeşil bayrak yarışın başlangıcını gösterir.	Yeşil bayrak yarışın başlangıcını gösterir.	Yeşil bir bayrak yarışın başlangıcını gösterir.
14	A <b>kid</b> in high school.	Lisede bir çocuk.	Lisede bir çocuk.	Lisede bir çocuk.	Lisedeki bir çocuk.
15	The children love to play in the <b>leaves</b> .	Çocuklar yapraklarda oynamayı severler.	Çocuklar yapraklarda oynamayı sever.	Çocuklar yapraklarda oynamayı severler.	Çocuklar yapraklarda oynamayı sever.
16	She was very <b>mean</b> to her sister.	Kız kardeşi için çok kaba davrandı.	Kız kardeşine çok kötü davranıyordu.	Kız kardeşine çok kötü davrandı.	Kız kardeşine çok kabaydı.
17	Every <b>move</b> was painful for the old man.	Yaşlı adam için her hareket acı vericiydi.	Her hareketi yaşlı adam için acı vericiydi.	Her hareket yaşlı adam için acı vericiydi.	Yaşlı adam için her hareket acı vericiydi.
18	A collection of precious <b>objects</b> .	Değerli nesnelere oluşan bir koleksiyon.	Değerli nesnelere oluşan bir koleksiyon.	Değerli eşyalardan oluşan bir koleksiyon.	Değerli eşyaların bir koleksiyonu.
19	Fill in the form on the <b>page</b> .	Sayfadaki formu doldurun.	Sayfadaki formu doldurun.	Sayfadaki formu doldurun.	Sayfadaki formu doldur.
20	A <b>pine</b> forest in the city.	Şehirde bir çam ormanı.	Şehirde bir çam ormanı.	Şehirde bir çam ormanı.	Şehirdeki bir çam ormanı.
21	At the <b>press</b> of a button.	Bir düğmeye basarak.	Bir düğmeye basarak.	Bir düğmeye basarak.	Bir düğmenin basınıda.

**Ek: 12 - Google, Bing, Yandex ve KMMÇ çıktıları tablosu (Madde 22 – 40)**

Madde	İNGİLİZCE CÜMLE	GOOGLE ÇIKTISI	BING ÇIKTISI	YANDEX ÇIKTISI	KMMÇ ÇIKTISI
22	The burning <b>desert</b> sun.	Yanan çöl güneşi.	Yanan çöl güneşi.	Yanan çöl güneşi.	Yanan çöl güneşi.
23	Build a <b>ship</b> for the navy.	Donanma için bir gemi inşa et.	Donanma için bir gemi inşa et.	Donanma için bir gemi inşa et.	Donanma için bir gemi inşa et.
24	I hate wearing a pink <b>tie</b> .	Pembe kravat takmaktan nefret ediyorum.	Pembe kravat takmak istemem.	Pembe kravat takmaktan nefret ediyorum.	Pembe bir bağ giymekten nefret ederim.
25	They don't know the <b>tie</b> between my grandmother and the town.	Büyükanemle kasaba arasındaki bağı bilmiyorlar.	Büyükanemle kasaba arasındaki bağı bilmiyorlar.	Büyükanemle kasaba arasındaki bağı bilmiyorlar.	Ninem ve şehir arasındaki bağı bilmezler.
26	Dinner will be a cold <b>buffet</b> .	Akşam yemeği soğuk açık büfe olacaktır.	Akşam yemeği soğuk bir büfe olacak.	Akşam yemeği soğuk büfe olacak.	Akşam yemeği bir soğuk büfe olacak.
27	My uncle and my father are very <b>close</b> .	Amcam ve babam çok yakınlar.	Amcam ve babam çok yakınlar.	Amcam ve babam çok yakınlar.	Amcam ve babam çok yakın.
28	The antique <b>console</b> will be in the living room.	Antik konsol oturma odasında olacak.	Antika konsol oturma odasında olacak.	Antika konsol oturma odasında olacak.	Antika konsol salonda olacak.
29	This is the new word <b>record</b> .	Bu yeni dünya rekoru.	Bu yeni dünya rekoru.	Bu yeni dünya rekoru.	Bu yeni dünya kayıttı.
30	I swatted the <b>fly</b> with a newspaper.	Sineği bir gazete ile karıştırdım.	Sineği bir gazeteyle ezdim.	Sineği bir gazete ile çırpıttım.	Bir gazete ile sineğe vurdum.
31	The leaves rustle in the <b>wind</b> .	Yapraklar rüzgarda hışırdar.	Yapraklar rüzgarda hışırdar.	Yapraklar rüzgarda hışırdıyor.	Yapraklar rüzgada hışırdar.
32	The <b>does</b> are walking in the forest.	Ormanda yürüyorlar.	Ormandakiler yürüyor.	Does ormanda yürüyor.	Ormanda geyikler yürümekte.
33	They sat on the grass in the <b>park</b> .	Parktaki çimlere oturdular.	Parkta çimlere oturdular.	Parkta çimlerin üzerine oturdular.	Parktaki çimende oturdular.
34	The gas stove was in the old <b>house</b> in the village.	Gaz sobası köydeki eski evdeydi.	Gaz sobası köydeki eski evdeydi.	Gaz sobası köydeki eski evdeydi.	Gaz sobası köydeki eski evdeydi.
35	They will be <b>live</b> on the city radio.	Şehir radyosundan canlı yayınlanacaklar.	Şehir radyosundan canlı yayınlanacaklar.	Şehir radyosunda canlı yayın yapacaklar.	Şehir radyosunda yayında olacaklar.
36	That person is really not your <b>type</b> .	Bu kişi gerçekten senin tipin değil.	O kişi gerçekten senin tipin değil.	O kişi senin tipin değil.	Gerçekten o kişi tipiniz değil.
37	My roommates ate two jars of <b>jam</b> in two days.	Oda arkadaşlarım iki günde iki kavanoz reçel yedi.	Oda arkadaşlarım iki günde iki kavanoz reçel yedi.	Oda arkadaşlarım iki günde iki kavanoz reçel yedi.	Oda arkadaşlarım iki günde iki kavanoz reçel yedi.
38	Endurance was the most important <b>attribute</b> in a teacher.	Dayanıklılık, bir öğretmende en önemli özellikti.	Dayanıklılık bir öğretmende en önemli özellikti.	Dayanıklılık bir öğretmenin en önemli özelliği idi.	Bir öğretilimde dayanıklılık en mühim vasıftı.
39	The old man's old <b>watch</b> is fast.	Yaşlı adamın eski saati hızlı.	Yaşlı adamın eski saati hızlı.	Yaşlı adamın eski saati çok hızlı.	Yaşlı adamın eski saati hızlı.
40	There is a huge <b>wave</b> of divorce in the developed countries.	Gelişmiş ülkelerde büyük bir boşanma dalgası var.	Gelişmiş ülkelerde büyük bir boşanma dalgası yaşanıyor.	Gelişmiş ülkelerde büyük bir boşanma dalgası var.	Gelişmiş ülkelerde büyük bir boşanma dalgası var.

**Ek: 13 - Google, Bing, Yandex ve KMMÇ çıktıları tablosu (Madde 41 – 61)**

Madde	İNGİLİZCE CÜMLE	GOOGLE ÇIKTISI	BING ÇIKTISI	YANDEX ÇIKTISI	KMMÇ ÇIKTISI
41	She <b>arms</b> herself with knowledge.	Kendini bilgiyle donatıyor.	Kendini bilgiyle silahlandırmak.	Kendini bilgiyle silahlandırıyor.	Bilgi ile kendini silahlandırır.
42	They <b>bear</b> the pain.	Acı çekiyorlar.	Acı çekiyorlar.	Acıyı taşıyorlar.	Acıya katlanırlar.
43	They are going to <b>beef</b> .	Sığır edecekler.	Sığır etiyle kavga ediyorlar.	Sığır eti yiyecekler.	Sızlanacaklar.
44	They <b>boil</b> the goat milk in a pan.	Keçi sütünü tavada kaynatırlar.	Keçi sütünü tavada kaynatıyorlar.	Keçi sütünü bir tavada kaynatırlar.	Bir tavadaki keçi sütünü kaynatırlar.
45	We <b>book</b> flights to big cities.	Büyük şehirlere uçuş rezervasyonu yapıyoruz.	Büyük şehirlere uçuş rezervasyonu yapıyoruz.	Büyük şehirlere uçuş rezervasyonu yapıyoruz.	Büyük şehirlere uçuşlar rezerve ederiz.
46	He is going to <b>box</b> in the Olympics in summer.	Yaz aylarında olimpiyatlarda boks yapacak.	Yaz aylarında olimpiyatlarda boks yapacak.	Yaz aylarında olimpiyatlarda boks yapacak.	Yazın Olimpiyatlarda dövüşecek.
47	The people <b>carp</b> about junk mail.	İnsanlar önemsiz posta hakkında konuşuyorlar.	İnsanlar önemsiz postalar hakkında sazani.	İnsanlar önemsiz posta hakkında sazani.	İnsanlar istenmeyen posta hakkında mızızlanırlar.
48	Her mom is going to <b>clip</b> her hair.	Annesi saçını kesecek.	Annesi saçını kesecek.	Annesi saçını kesecek.	Annesi saçını kesecek.
49	They <b>crane</b> forward.	İleriye doğru döner.	Öne doğru vinçle ilerliyorlar.	Vinç ileri.	İleriye boynunu uzatırlar.
50	I'd like to <b>express</b> my gratitude to everyone.	Herkese şükranlarımı sunmak isterim.	Herkese şükranlarımı sunuyorum.	Herkese şükranlarımı sunmak istiyorum.	Herkese minnettarlığımı ifade etmek isterim.
51	She is going to <b>drop</b> the book.	Kitabı bırakacak.	Kitabı bırakacak.	Kitabı bırakacak.	Kitabı düşürecek.
52	Muslims <b>fast</b> during Ramadan.	Müslümanlar Ramazan ayında oruç tutar.	Müslümanlar Ramazan ayında oruç tutarlar.	Ramazan'da oruç Müslümanlar.	Ramazan süresince müslümanlar oruçlu tutar.
53	<b>Flag</b> the potential problems in our proposal.	Önerimizdeki olası sorunları işaretleyin.	Teklifimizdeki olası sorunları işaretle.	Teklifimizdeki olası sorunları işaretleyin.	Teklifimizdeki muhtemel problemleri işaretle.
54	Don't <b>kid</b> yourself.	Kendinizi kandırmayın.	Kendini kandırma.	Kendini kandırma.	Kendini kandırma.
55	Their father <b>leaves</b> home for work.	Babaları iş için evden ayrılıyor.	Babaları işe giderken evden ayrılır.	Babaları iş için evden ayrılıyor.	İş için babaları evden ayrılır.
56	I didn't <b>mean</b> any harm.	Zarar vermek istemedim.	Kötü bir şey demek istemedim.	Kötü bir niyetim yoktu.	Hiç zararı istemedim.
57	They <b>move</b> the flag up and down.	Bayrağı yukarı ve aşağı hareket ettirirler.	Bayrağı yukarı ve aşağı hareket ettiriyorlar.	Bayrağı yukarı ve aşağı hareket ettirirler.	Yukarı aşağı bayrağı taşırlar.
58	People at my work <b>object</b> to the book.	İnsanlar kitaba itiraz ediyor.	İnsanlar kitaba itiraz etti.	İnsanlar kitaba itiraz ediyor.	İnsanlar kitaba itiraz eder.
59	<b>Page</b> the man in the hospital.	Hastanedeki doktora çağrı yapın.	Hastanedeki doktora çağrıla.	Hastanede doktor Page.	Hastanedeki doktora çağrı.
60	they still <b>pine</b> for their lost wealth.	Hala kayıp servetleri için özlemlerini sürdürüyorlar.	Hala kayıp servetleri için tınıyorlar.	Hala kayıp servetleri için can atıyorlar.	Kayıp servetleri için hala özlem çekerekler.
61	<b>Press</b> the button for the third floor.	Üçüncü katın düğmesine basın.	Üçüncü kat için düğmeye basın.	Üçüncü kat için düğmeye basın.	Üçüncü kat için düğmeye bas.

**Ek: 14 - Google, Bing, Yandex ve KMMÇ çıktıları tablosu (Madde 62 – 80)**

Madde	İNGİLİZCE CÜMLE	GOOGLE ÇIKTISI	BING ÇIKTISI	YANDEX ÇIKTISI	KMMÇ ÇIKTISI
62	He <b>deserted</b> his wife and children for another woman.	Karısını ve çocuklarını başka bir kadın için terk etti.	Karısını ve çocuklarını başka bir kadın için terk etti.	Başka bir kadın için karısını ve çocuklarını terk etti.	Başka bir kadını için hanım ve çocuklarını terk etti.
63	The company <b>ships</b> its goods to twenty countries.	Şirket mallarını yirmi ülkeye gönderiyor.	Şirket mallarını yirmi ülkeye gemi iletir.	Şirket mallarını yirmi ülkeye gönderiyor.	Şirket yirmi ülkeye malını yollar.
64	The robbers were going to <b>tie</b> the men to a pipe with cable.	Soyguncular, adamları kabloyla bir boruya bağlayacaklardı.	Soyguncular adamları kabloyla boruya bağlayacaklardı.	Soyguncular adamları kabloyla bir boruya bağlayacaklardı.	Hırsızlar kablo ile bir boruya adamları bağlayacaktı.
65	The buildings <b>house</b> the young students.	Binalar genç öğrencileri barındırıyor.	Binalar genç öğrencileri barındırıyor.	Binalar genç öğrencilere ev sahipliği yapıyor.	Binalar genç öğrencileri barındırır.
66	The waves are going to <b>buffet</b> the shore.	Dalgalar kıyıya vuracak.	Dalgalar kıyıya açık büfe olacak.	Dalgalar kıyıya açık büfe olacak.	Dalgalar sahili dövecek.
67	We <b>close</b> the office at noon on Fridays.	Cuma günleri öğle saatlerinde ofisi kapatıyoruz.	Cuma günleri öğlen ofisi kapatıyoruz.	Cuma günleri öğlen ofisini kapatıyoruz.	Cuma günleri öğle vakti ofisi kapatırız.
68	She is going to <b>console</b> herself with the cat.	Kendini kediyile teselli edecek.	Kediyile kendini avutacak.	O kedi ile kendini teselli edecek.	Kedi ile kendini avutacak.
69	These cameras <b>record</b> 25 frames per second.	Bu kameralar saniyede 25 kare kaydeder.	Bu kameralar saniyede 25 kare kaydeder.	Bu kameralar saniyede 25 kare kaydeder.	Bu kameralar saniyede 25 kareyi kayda alır.
70	He was the first person to <b>fly</b> .	Uçan ilk kişiydi.	Uçan ilk kişiydi.	İlk uçan kişi oydu.	ilk uçan kişiydi.
71	The little girls were going to <b>fly</b> the kite.	Küçük kızlar uçurtmayı uçuracaklardı.	Küçük kızlar uçurtmayı uçuracaktı.	Küçük kızlar uçurtmayı uçuracaktı.	Küçük kızlar uçurtmayı uçuracaktı.
72	My daughter <b>does</b> her homework every night.	Kızım her gece ödevini yapıyor.	Kızım her gece ödevini yapıyor.	Kızım her gece ödevini yapıyor.	Her gece kızım ev ödevini yapar.
73	They <b>park</b> their cars in the building.	Arabaları binaya park ediyorlar.	Arabaları binaya park ediyorlar.	Arabaları binaya park ediyorlar.	Binadaki arabaları park ederler.
74	We are going to <b>park</b> in front of the TV.	Televizyonun önüne park edeceğiz.	Televizyonun önüne park edeceğiz.	Televizyonun önüne park edeceğiz.	Televizyon önünde park edeceğiz.
75	He wants to <b>live</b> the last years of his life with his dogs.	Hayatının son yıllarını köpekleriyle yaşamak istiyor.	Hayatının son yıllarını köpekleriyle geçirmek istiyor.	Hayatının son yıllarını köpekleriyle yaşamak istiyor.	Köpekleri ile hayatının son yıllarını yaşamak ister.
76	I <b>type</b> over 100 words per minute.	Dakikada 100 kelimedenden fazla yazıyorum.	Dakikada 100'den fazla kelime yazıyorum.	Dakikada 100'den fazla kelime yazıyorum.	Dakikada 100 kelime üzerinde yazırım.
77	Their man's gun will <b>jam</b> .	Adamlarının silahı sıkışacak.	Adamlarının silahı çalacak.	Adamlarının silahı sıkışacak.	Adamlarlarının silahı sıkışacak.
78	She will <b>attribute</b> her success to her late father's hard work.	Başarısını rahmetli babasının sıkı çalışmasına bağlayacak.	Başarısını rahmetli babasının sıkı çalışmasına bağlayacak.	Başarısını geç babasının sıkı çalışmasına bağlayacak.	Ölen babasının çok çalışmasına başarısını dayandıracak.
79	Our kids only <b>watch</b> YouTube videos at weekends.	Çocuklarımız YouTube videolarını sadece hafta sonları izliyor.	Çocuklarımız YouTube videolarını sadece hafta sonları izliyor.	Çocuklarımız sadece hafta sonları YouTube videolarını izliyor.	Hafta sonları çocuklarımız sadece YouTube videolarını izler.
80	The people on the bus <b>wave</b> to the villagers.	Otobüsteki insanlar köylülere el sallıyor.	Otobüstekiler köylülere el sallıyor.	Otobüsteki insanlar köylülere el sallıyor.	Otobüsdeki insanlar köylülere el sallıyor.

## ÖZGEÇMİŞ

### EĞİTİM:

1973 yılında İstanbul'da doğdu. İlkokul, ortaokul ve liseyi İstanbul'da bitirdi. Lisans eğitimini 1993 – 1997 yılları arasında Marmara Üniversitesi, İngiliz Dili Eğitimi Bölümü'nde tamamladı. 2015 yılında Bahçeşehir Üniversitesi, Bilgi Teknolojileri, Yüksek Lisans programına kaydoldu ve mezun olduğu 2017 yılında İstanbul Sabahattin Zaim Üniversitesi, Bilgisayar Bilimleri ve Bilgisayar Mühendisliği doktora bölümüne kabul edildi.

### MESLEKİ DENEYİM:

1986 – 1990 yılları arasında, çıraklıkla başladığı ve devlet meslek edindirme kursu sonrası devam ettirdiği çalışma hayatında, elektronik teknisyenliği yaptı. Lisans mezuniyeti sonrası, devlet okullarında İngilizce öğretmenliği yapmaya başladığı 2006 yılına kadar çeşitli özel eğitim kurumlarında görev yaptı.