

T.C.
SAKARYA ÜNİVERSİTESİ
BİLGİSAYAR VE BİLİŞİM BİLİMLERİ FAKÜLTESİ

BSM 498 BİTİRME ÇALIŞMASI

TÜRKÇE DİLİNİN MORFOLOJİK
İNCELEMESİ VE BİÇİMSEL OLARAK
TANIMLANMASI

B121210057 – Egemen ZEYTİNCİ

G101210009 – Taşkın TEMİZ

Fakülte Anabilim Dalı : BİLGİSAYAR MÜHENDİSLİĞİ
Tez Danışmanı : Prof. Dr. Nejat YUMUŞAK

2015-2016 Bahar Dönemi

T.C.
SAKARYA ÜNİVERSİTESİ
BİLGİSAYAR VE BİLİŞİM BİLİMLERİ FAKÜLTESİ

**TÜRKÇE DİLİNİN MORFOLOJİK
İNCELEMESİ VE BİÇİMSEL OLARAK
TANIMLANMASI**

BSM 498 - BİTİRME ÇALIŞMASI

**Egemen ZEYTİNCİ
Taşkın TEMİZ**

Fakülte Anabilim Dalı : BİLGİSAYAR MÜHENDİSLİĞİ

Bu tez .. / .. / ... tarihinde aşağıdaki jüri tarafından oybirliği / oyçokluğu ile kabul edilmiştir.

.....
Jüri Başkanı

.....
Üye

.....
Üye

ÖNSÖZ

Bitirme çalışmamız sırasında desteğini ve ilgisini eksik etmeyen hocamız Prof. Dr. Nejat Yumuşak’a teşekkürü borç biliriz. Ayrıca eğitim hayatımız boyunca bize her daim destek olan ailelerimize çok teşekkür ederiz.

İÇİNDEKİLER

ÖNSÖZ.....	iii
İÇİNDEKİLER.....	iv
SİMGELER VE KISALTMALAR LİSTESİ.....	vi
ŞEKİLLER LİSTESİ.....	vii
TABLolar LİSTESİ.....	ix
ÖZET.....	x
 BÖLÜM 1.	
GİRİŞ.....	1
1.1. Derleyicilerin	2
Tarihçesi.....	
1.2. Derleyici Aşamaları.....	2
 BÖLÜM 2.	
TÜRKÇE’NİN BİÇİMBİLİMSEL ÇÖZÜMLEMESİ.....	6
2.1. Türkçe’nin Yapısı.....	6
2.1.1. Çekim Kümeleri	6
2.1.2. Eklerin Diziliş Kuralları.....	10
2.1.3. İsim Çekim Ekleri Kümesi.....	12
2.1.4. Ek-Eylem Ekleri Kümesi.....	18
2.1.5. Eylem Zaman Ekleri Kümesi.....	20
2.1.6. Eylem Çekim Ekleri Kümesi.....	26
2.1.7. Yapım Ekleri Kümesi.....	34
 BÖLÜM 3.	
SONUÇLAR VE ÖNERİLER.....	36
 KAYNAKLAR.....	37
EK A.....	38
EK B.....	39
ÖZGEÇMİŞ.....	40
 BSM 498 BİTİRME ÇALIŞMASI DEĞERLENDİRME VE SÖZLÜ SINAV	
TUTANAĞI.....	41

SİMGELER VE KISALTMALAR LİSTESİ

()	: içerisindeki harf ek içinde yer almayabilir
A	: a veya e harfi yerine
C	: c veya ç harfi yerine
ÇK	: Çekim kümeleri
D	: d veya t harfi yerine
H	: ı, i, u veya ü harfi yerine
I	: ı veya i harfi yerine
M	: Master
SDM	: Sonlu durum makinesi
TS	: Türetim sınırı

ŞEKİLLER LİSTESİ

Şekil 1.1.	Sözcüksel analiz örneği.....	3
Şekil 1.2.	Orta düzey kod üretimi.....	5
Şekil 1.3.	Kod optimizasyonu.....	5
Şekil 1.4.	Derleyici işleyiş diyagramı.....	5
Şekil 2.1.	Çekim ekleri ve türetim sınırları.....	7
Şekil 2.2.	Biçimbilimsel Çözümleyici ile Çözümleme.....	7
Şekil 2.3.	Türetim sınırları.....	7
Şekil 2.4.	Bağlılık ilişkisi.....	8
Şekil 2.5.	Cümlecğin çekim eklerine ayrılması.....	9
Şekil 2.6.	Türkçe cümlelerde cümle içi bağlılıklar.....	10
Şekil 2.7.	İsim çekim ekleri soldan sağa sonlu durum makinesi.....	13
Şekil 2.8.	"Sandıktakilerden" sözcüğünün soldan sağa incelemesi.....	15
Şekil 2.9.	İsim çekim ekleri sağdan sola sonlu durum makinesi.....	16
Şekil 2.10.	"Etkilerden" sözcüğünün sağdan sola incelenmesi.....	16
Şekil 2.11.	"Arabamız" sözcüğünün sağdan sola incelenmesi.....	17
Şekil 2.12.	Birden çok biçimbilimsel analiz gerektiren sözcük örneği...	17
Şekil 2.13.	"Akıllıymışsın" ve "okuldaysalar" sözcüklerinin soldan sağa SDM ile incelenmesi.....	19
Şekil 2.14.	Ek-eylem ekleri soldan sağa sonlu durum makinesi.....	19
Şekil 2.15.	Ek-eylem ekleri sağdan sola sonlu durum makinesi.....	20
Şekil 2.16.	Sözcüklerin ek-eylem ekleri sağdan sola SDM ile incelenmesi.....	20
Şekil 2.17.	"Görüyorsunuz" sözcüğünün incelemesi.....	22
Şekil 2.18.	"Yapmasanız" sözcüğünün incelemesi.....	22
Şekil 2.19.	"Yapmazlarsa" sözcüğünün incelemesi.....	22
Şekil 2.20.	Eylem zaman ekleri soldan sağa sonlu durum makinesi.....	23
Şekil 2.21.	Eylem zaman ekleri sağdan sola sonlu durum makinesi.....	24
Şekil 2.22.	"Okudum" sözcüğünün incelemesi.....	24

Şekil 2.23.	"Yapmasanız" sözcüğünün incelemesi.....	25
Şekil 2.24.	"Yapmazlarsa" sözcüğünün incelemesi.....	25
Şekil 2.25.	"Gezermişiz" sözcüğünün incelemesi.....	26
Şekil 2.26.	Eylem çekim ekleri soldan sağa sonlu durum makinesi.....	29
Şekil 2.27.	"Görüşmezler" sözcüğünün incelemesi.....	29
Şekil 2.28.	"Gönderilebilir" sözcüğünün incelemesi.....	30
Şekil 2.29.	"Okuyucular" sözcüğünün incelemesi.....	30
Şekil 2.30.	"Okumak" sözcüğünün incelemesi.....	30
Şekil 2.31.	"Okumaktadır" sözcüğünün incelenmesi.....	30
Şekil 2.32.	"Okuyamazmışım" sözcüğünün incelemesi.....	31
Şekil 2.33.	Eylem çekim ekleri sağdan sola sonlu durum makinesi.....	31
Şekil 2.34.	Sonlu durum makinesi girişleri.....	31
Şekil 2.35.	"Görüşmezler" sözcüğünün incelemesi.....	32
Şekil 2.36.	"Gönderebilir" sözcüğünün incelemesi.....	33
Şekil 2.37.	"Okuyamazmışım" sözcüğünün incelemesi.....	33
Şekil 2.38.	"Okuyucuların" sözcüğünün incelemesi.....	33
Şekil 2.39.	"Yürümektendir" sözcüğünün incelemesi.....	34

TABLÖLAR LİSTESİ

Tablo 2.1	İsim çekim ekleri ek listesi.....	12
Tablo 2.2.	İsim çekim ekleri ek listesi.....	13
Tablo 2.3.	Ek-eylem ekleri ek listesi.....	18
Tablo 2.4.	Eylem zaman ekleri ek listesi.....	21
Tablo 2.5.	Eylem çekim ekleri ek listesi.....	27
Tablo 2.6.	Eylem çekim ekleri ek listesi.....	28
Tablo 2.7.	Yapım ekleri ek listesi.....	35

ÖZET

Anahtar kelimeler: Türkçe'nin Morfolojik İncelemesi, Derleyici Aşamaları, Türkçe'nin Biçimsel Tanımı, Biçimsel Makineler

Türkçe dilinin morfolojik olarak incelenmesi ve biçimsel makinelerle tanımlanması, günümüz şartlarında önemli bir noktaya temas etmektedir. Yüksek başarı oranlarının yakalandığı; fakat gerçekleştirme aşamasında tam bir form elde edilememiş olan bu alan, başta Türkçe olmak üzere sondan eklemeli dil ailelerine mensup diğer diller için de önem arz etmektedir. Özellikle Türkçe gibi durmaksızın yenilenen ve neredeyse sonsuz kelime türetme olanağına sahip dillerin biçimsel aşamada gerçekleştirilmesi arama motorları, çeviri kanalları gibi birçok alan için oldukça önemlidir.

Türkçe dilinin, morfolojik incelemeler esnasında ele alınan sözcüksel, sözdizimsel ve anlamsal analizinden faydalanarak derleyici aşamaları gerçekleştirilmiştir. Bu çalışmada, Türkçe gibi sondan eklemeli dillerin biçimbilimsel olarak tanımlanması ve günümüzde bu teknolojiye ihtiyaç duyan arama motorları, çeviri kanalları, sözcük işlem programları gibi alanlarda Türkçe'nin daha anlaşılır hale gelmesi hedeflenmiştir. Nitekim çalışmalar esnasında elde edilen bilgilerin ve deneyimin ışığında Türkçe'nin morfolojik incelenmesi tamamlanmış, biçimsel olarak tanımlanması sağlanmıştır.

BÖLÜM 1. GİRİŞ

Herhangi bir dilde programlama yapmak için öncelikle o dile özel bir programa ihtiyaç duyarız. Bu program kaynak kodumuzu yazabileceğimiz ve kodumuzu çalıştırabileceğimiz bir yapıya sahiptir. İşte bu özel programa “derleyici” adı verilir.

Derleyici kaynak dildeki bir programı hedef bir dile ya da hedef bir makineye uygun dile dönüştürür. Makine dilinde programlama yavaş, hataya açık ve zordur. Kullanılan bilgisayarın mimarisi ile ilgili bilgi gerektirir. Ayrıca makine değiştiğinde kodu kullanmak mümkün olmaz. Derleyici, programımızı üst düzeyde rahat bir şekilde yazarak istediğimiz bilgisayarda çalıştırmamızı sağlar. Bizi makine düzeyine inmekten kurtarır. Derleyici kaynak olarak aldığı kodu çalıştırabilir koda dönüştürür.

Derleyici, kaynak kodumuzu kolay bir şekilde yazmak, hızlıca gözden geçirebilmek, yazım hatalarımızı daha kolay görebilmek gibi konularda, bize yardımcı olmak amacıyla gelişmiş bir metin düzenleyicisine sahiptir. Tabii ki, bu özellik günümüz derleyicileri için geçerlidir. Eski yıllardaki derleyiciler bu konuda daha zayıftılar.

Yazdığımız kodu derlediğimizde derleyici, varsa mantık ya da söz dizimi hatalarımızı listeleyecektir. Ayrıca çalışma sırasında sorun çıkarabilecek ya da fark edilemeyecek farklı çalışmalara sebep olabilecek durumları da uyarı olarak listeler.

Derleme işlemini belirli satırlara kadar yapmamız ya da her satırın etkilerini adım adım gözlemlememiz de mümkündür. Bunların dışında derleyicilerin geliştirmeyi kolaylaştırıcı birçok özelliği olabilir. Bu özellikler derleyiciyi geliştirene göre değişir. Bazı dillerin ilk oluşturuldukları zamanlarda tek derleyicisi vardır. Bu diller genelde bir yazılım firması tarafından yazılmıştır ve derleyicileri de bu firma tarafından oluşturulur. Dil, bu ilk derleyicisiyle tanınır [1].

1.1. Derleyicilerin Tarihçesi

Bilgisayar yazılımları uzun yıllar boyu çevirici dil (assembly language) sayesinde yazılmıştı. Bir derleyici yazmanın maliyetinin bir yazılımı farklı işlemci türlerinde kullanabilmenin maliyetinden daha düşük olduğunun farkedilmesi ile de, yüksek seviyeli diller icat edilmiş ve ilk derleyiciler ortaya çıkmaya başlamıştır. İlk derleyici de Amerikalı deniz subayı ve bilgisayar bilimcisi Grace Hopper tarafından 1952 yılında A-0 programlama dili için yazılmıştır. Bunun dışında FORTRAN ve COBOL gibi diller de farklı mimarilerde derlenebilen diller olarak ortaya çıkmışlardır. Ve bu derleyiciler çevirici dilde yazılmışlardır.

Derleyici yapımı ve optimizasyonu günümüzde bilgisayar mühendisliği ve bilgisayar bilimi bölümlerinde öğretilmektedir. Bu derslerde genellikle eğitim amaçlı bir programlama dilinin derleyicisinin de uygulanması yapılmaktadır.

Derleyiciler bugün kullandığımız makineden bağımsız yazılımların geliştirilmesini sağlamışlardır. İlk yüksek seviyeli dil olan FORTRAN'ın geliştirilmesinden önce makine bağımlı çevirici dil (assembly language) yaygın bir şekilde kullanılmaktaydı. Çevirici dil aynı mimaride daha çok yeniden kullanılabilen ve yer değiştirebilen program oluşturmaya rağmen bu programı başka bir donanım mimarisinde çalıştırmak için programı modifiye etmek veya baştan yeniden yazmak gerekliydi. Fakat FORTRANDAN sonra COBOL, C, BASIC yüksek seviyeli programlama dillerinin ortaya çıkmasıyla birlikte bilgisayar programcıları makineden bağımsız programlar yazabilmeye başladılar [2].

1.2. Derleyici Aşamaları

Derleyici yukarıda da bahsettiğimiz gibi yüksek seviyeli dilde yazılan programı makina koduna çevirmekle sorumludur. Derleyiciler üç temel kısımdan oluşur:

Ön-son (Frontend): Programın yazılan dilin kurallarına ve semantiklerine uygun bir şekilde yazılıp yazılmadığını kontrol eder. Burada dilin kurallarına uyan ve uymayan programlar belli olur, varsa hatalar rapor edilir, uyarılar verilir. Ön-son , orta-son için IR(intermediate representation) oluşturur.

Orta-son (Middle-end): Bu kısım optimizasyonun yapıldığı yerdir. Yararsız ve ulaşılamayan kodlar silinir, sabit değerler bulunur. Orta-son da arka-son için IR oluşturur.

Arka-son (Backend): Orta-sondan gelen IR'yi istenen çeviri koduna dönüştürür. Değişkenler de registerlar için burada seçilir [2].

Bu temel kısımlar da birtakım alt başlıklara ayrılır. Örneğin Ön-son (Frontend) kısmı üç başlık altında incelenebilir.

Sözcüksel Analiz: Program metninin okunduğu ve analiz edildiği ilk kısımdır. Kaynak kod derleyiciye ilk geldiği zaman, ilk olarak, kaynak kodun harfleri teker teker işlenir ve tanımlanarak “token” adlı nesnelere dönüştürülür. Token; değişken adı, anahtar kelime ya da sayı gibi programlama dilindeki bir sembolü gösterir. Örnek olarak verilen kod bloğunun analizi Şekil 1.1 de incelenebilir.

<code>newval:=oldval+12</code>	tokenler:	<code>newval</code>	değişken
		<code>:=</code>	atama operatörü
		<code>oldval</code>	değişken
		<code>+</code>	toplama operatörü
		<code>12</code>	sayı

Şekil 1.1. Sözcüksel analiz örneği

Sözcüksel analizi gerçekleştiren derleyici parçasına “lexer” denir. Lexerin görevi, yazım analizi yapıp sonuç olarak bir token listesi oluşturmaktır. Tokenleri tanımlamak için düzgün ifadeler (regular expressions) kullanılır. Sözcüksel analizin uygulanmasında sonlu durum otomatu (finite state automata) kullanılabilir. Bu yazım analizi sırasında kaynak koddaki hatalar da bulunabilir. Bu hatalar lexer tarafından tespit edilir ve yazım hatası (syntax error) olarak bildirilir. Lexer, kaynak kodun ifade ettiği anlam ile uğraşmaz. Hangi tokenin hangisinden sonra geldiği ile ilgilenmez [1].

Sentaktik (Sözdizimsel) Analiz: Parser, lexerın oluşturduğu token listesini alır ve bir ağaç yapısı üretir. Bu oluşturulan ağaca “parse ağacı” denir. Token listesinden oluşturulmuş olan parse ağacı yine tokenlerden oluşmuş bir yapıdır ve derleyicinin

tanıldığı gramere göre oluşturulur. Parser, ağacı oluştururken gramere uygun olmayan bir token ile karşılaşırsa bunu hata olarak bildirir. Parserın ürettiği bu hataya “gramer hatası” ya da “parse hatası (parse error)” denir.

Parser, token içinde bulunan ve tokenin ifade ettiği metin ile uğraşmaz. Tokenin karakter katarında hangi pozisyonda bulunduğu ile ilgilenmez. Lexer gibi tokenin ifade ettiği anlam ile uğraşmaz. Parser, yalnızca token listesinin dizilişine göre gramere uygunluğu kontrol eder. Bunun sonucunda da bu token listesinden oluşan bir parse ağacı üretir [3].

Semantik (Anlamsal) Analiz: Parse ağacı oluştuktan sonra bu ağaç, semantik analize (semantic analysis) tabi tutulur. Ağacın her yaprağında bulunan tokenler, bu aşamada derleyici için bir anlam ifade etmeye başlar. Örneğin; derleyici tarafından integer olan bir tokene integer, değişken olan tokene ise değişken olarak anlam verilir. Bunun sonucunda anlam olarak belli hatalar oluşabilir. Bu aşamada oluşabilecek hatalara “semantik hatalar (semantic errors)” denir. Örneğin; integer olan bir değişken, string olan bir değişkene direkt olarak eşitlenemez çünkü bunların token tipleri farklıdır.

Bu aşamada, kaynak programdaki anlamsal hatalar kontrol edilir ve kod üretimi için veri tipi bilgileri belirlenir. Tip kontrolü anlamsal analizin en önemli kısmıdır [3].

Orta-son (Middleend) ise Orta Düzey Kod Üretimini içinde barındırır. Semantik analiz sonucunda herhangi bir hata oluşmamış ise parse ağacı yardımı ile “intermediate code” ya da “intermediate representation” denilen ara bir kod oluşturulur. Bu ara kod oluşturulduktan sonra kodlar optimize edilir ve tekrar son çıktı için kod üretilir. Bu safha her derleyicide bulunmayabilir. Bazı derleyiciler ara kod oluşturmada direkt olarak kod oluşturabilir. Orta Düzey Kod Üretiminin çalışması Şekil 1.2 de gösterilmiştir [3].

```
id1 := id2 * id3 + 1
```

```
MULT id2, id3, temp1
```

```
ADD temp1, #1, temp2
```

```
MOV temp2, id1
```

Şekil 1.2. Orta düzey kod üretimi

Arka-son (Backend) kısmı iki başlık altında incelenebilir. Bu kısımlar Kod Optimizasyonu ve Kod Üretimi olarak adlandırılır.

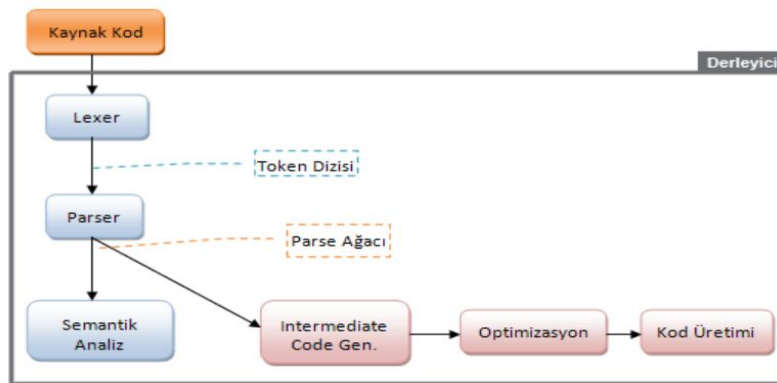
Kod Optimizasyonu: Programı tam otomatik olarak değiştirerek; daha hızlı çalışmasını, daha az bellek kaplamasını ve genel olarak kaynak kullanımında tutumlu davranmayı sağlamaktadır. Kod Optimizasyonunun çalışması sonrasında Şekil 1.2 deki kodun son hali Şekil 1.3. de gösterilmiştir.

```
MULT id2, id3, temp1
ADD temp1, #1, id1
```

Şekil 1.3. Kod optimizasyonu

Kod Üretimi: İstenen mimaride hedef dili üretebilir. Genelde assembly kodu üretir. Bu daha sonra assembler (çevirici, toplayıcı) tarafından çalıştırılabilir koda çevrilir. Başka bir dile çeviri şeklinde de olabilir. Bu, insanların yaptığı çeviriyle özdeşleştirilebilir [3].

Derleyici işleyiş diyagramı Şekil 1.4 de verilmiştir [3].



Şekil 1.4. Derleyici işleyiş diyagramı

BÖLÜM 2. TÜRKÇE’NİN BİÇİMBİLİMSEL ÇÖZÜMLEMESİ

Altay dil ailesinde yer alan ve bitişken bir dil olan Türkçe, biçimbilimsel olarak oldukça karmaşık bir yapıya sahiptir. Bu bölümde öncelikle Türkçe’nin bilişim biçimbilimsel çözümlemesi üzerine yapılan araştırmalardan ve uygulamalardan kısaca bahsedilecek, daha sonra da Türkçe’nin biçimbilimsel yapısı ve eklerin art arda sıralanış kuralları (morfolotik kuralları) ile ilgili ayrıntılı bilgi verilecektir [4].

2.1. Türkçe’nin Yapısı

Bitişken bir dil olan Türkçe çok zengin biçimbilimsel bir yapıya sahiptir. Sözcükler sonlarına art arda ekler konularak yüzlerce farklı sözcüğe dönüştürülebilirler. Birçok dilde sözcükten ayrı olarak yazılan ilgeçler, Türkçede genelde bir sonek olarak kelimeye eklenip tek bir sözcük oluştururlar. Benzer şekilde kişi, yardımcı eylem gibi birçok ayrı yazılan sözcük, Türkçede yine ekler vasıtasıyla ifade edilirler. Bu nedenle, bir başka dilde uzun bir cümle ile ifade edilen deyişlerin Türkçe’de tek bir sözcük ile ifade edilmesi çok rastlanan bir durumdur. Türkçe’de bir sözcüğün ekler yardımı ile dönüştürülebileceği farklı sözcük sayısı kuramsal olarak sonsuzdur [4].

2.1.1. Çekim Kümeleri

Türkçede sözcükler sonlarına eklenen eklerle farklı türde sözcüklere de dönüşebilirler; eylemler isimlere, isimler eylemlere vb. Türkçe’nin bu özelliği ilgili yayınlarda (Ofłazer, Say, Tür, & Tür, 2003), (Ofłazer, 2003), (Eryiğit, Statistical Dependency Parsing of Turkish, 2006), (Eryiğit, 2006), (Hakkani-Tür, Ofłazer, & Tür, 2002) sözcüklerin çekim kümelerine (ÇK) ayrılması biçiminde gösterilmektedir. Bu gösterimde, Türkçe bir sözcüğün bir dizi çekim kümesinden oluştuğı ve bu ÇK’lerinin türetim sınırlarından (TS) bölündüğü varsayılmaktadır. Bu özellik aşağıdaki gibi gösterilmektedir [6] [7] [8]:

$$\text{gövde} + \text{ÇK}_1 + \text{^TS} + \text{ÇK}_2 + \text{^TS} + \dots + \text{^TS} + \text{ÇK}_n.$$

Şekil 2.5. Çekim ekleri ve türetim sınırları

Burada her ÇKi , ilgili olduğu çekim kümesine ait biçimbilimsel özellikleri ve sözcük sınıflarını belirtmektedir. Aşağıda, bu durum türemiş bir niteleyici olan ``sağlamlaştırdığımızdaki" sözcüğünün Oflazer (1994)'in biçimbilimsel çözümleyicisi ile çözümlemesi üzerinde gösterilmektedir. Bu gösterimde kullanılan etiketler şöyledir: +Adj: Sıfat, +Verb: Eylem, +Become: oluşmak, +Caus: Ettirgen, +Pos: olumlu, +Noun: İsim, +PastPart: geçmiş zaman ortacı, +A3sg: 3. tekil kişi kişi/sayı uyum imi, +P3sg: 3. tekil kişi iyelik imi, +Loc: -de hali, +Rel: ilişkilendirici [8].

sağlam+Adj

+^TS+Verb+Become

+^TS+Verb+Caus+Pos

+^TS+Noun+PastPart+A3sg+Plpl+Loc

+^TS+Adj+Rel

Şekil 2.6 Biçimbilimsel Çözümleyici ile Çözümleme

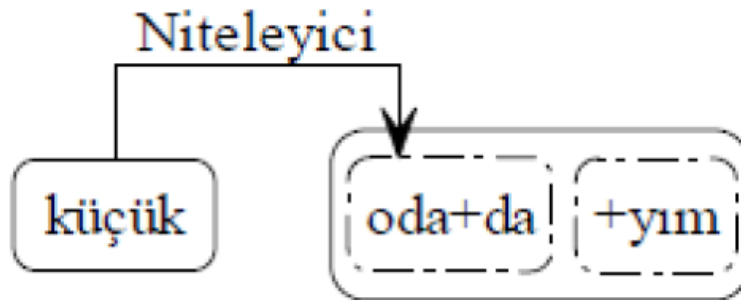
^TS sınırları sözcük üzerinde gösterilmek istenirse şöyle görünecektir [4]:

| sağlam | ^TS | laş | ^TS | tır | ^TS | dığımızda | ^TS | ki |
 ÇK₁ ÇK₂ ÇK₃ ÇK₄ ÇK₅

Şekil 2.7. Türetim sınırları

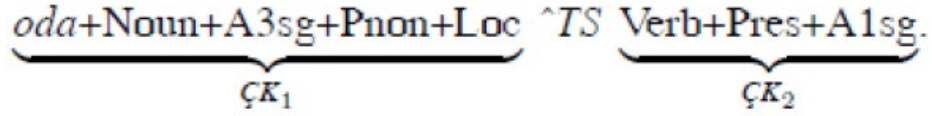
Buradaki beş çekim kümesi, \wedge TS türetim sınırı işaretleri ile birbirinden ayrılmış özellik dizileridir. İlk ÇK gövdenin tek özelliği olan sözcük sınıfını göstermektedir. “sağlam” sözcüğü bir sıfattır. İkinci ÇK, önceki sığata “oluşmak” anlamı katılarak bir eylem türetmeyi göstermektedir. Üçüncü ÇK, önceki eylemden olumlu bir ettirgen eylemin türetildiğini belirtmektedir. Dördüncü ÇK, alt sözcük sınıfı olarak geçmiş zaman ortacı taşıyan, birinci çoğul kişi iyelik ve -de hal eki almış bir isimin türetilmesini belirtmektedir. Son olarak da, beşinci ÇK, ilişkilendirici bir sıfat türetilmesini belirtmektedir [4].

Türkçe’nin bilişimsel dil işlemeş konusunda çalışanların sözcüklerden daha küçük birimler (ÇK’leri) tanımlama ihtiyaçları, Türkçe bir cümledeki cümle içi ilişkilerin sözcüklerin alt bölümleri arasında gerçekleşebilmesinden doğmaktadır. Türkçe’nin türetim sistemi çok üretkendir ve bir sözcüğün uydu veya iye olarak içerisinde bulunduğı cümle yapısı ilişkileri, sözcüğün içerdigi bir veya daha fazla türemiş yapının biçimbilimsel özellikleriyle belirlenmektedir. Bu ilişkileri, sadece sözcükler arasında göstermek çoğı doğal dil işleme uygulaması için yeterince anlamlı bilgi taşımamaktadır. Bu durum Şekil 2.4’te “Küçük odadayım” cümleciğı içerisindeki bağıllık ilişkisi ile örneklenmektedir. Şekildeki gösterimde sözcükler dikdörtgenler içerisinde ve bunların içerdikleri çekim kümeleri de noktalı dikdörtgenler içerisinde gösterilmektedirler [4].



Şekil 2.8. Bağıllık ilişkisi

Şekilde 2.5'teki “odadayım” sözcüğü iki ÇK'den oluşmaktadır:



Şekil 2.9. Cümlecğin çekim eklerine ayrılması

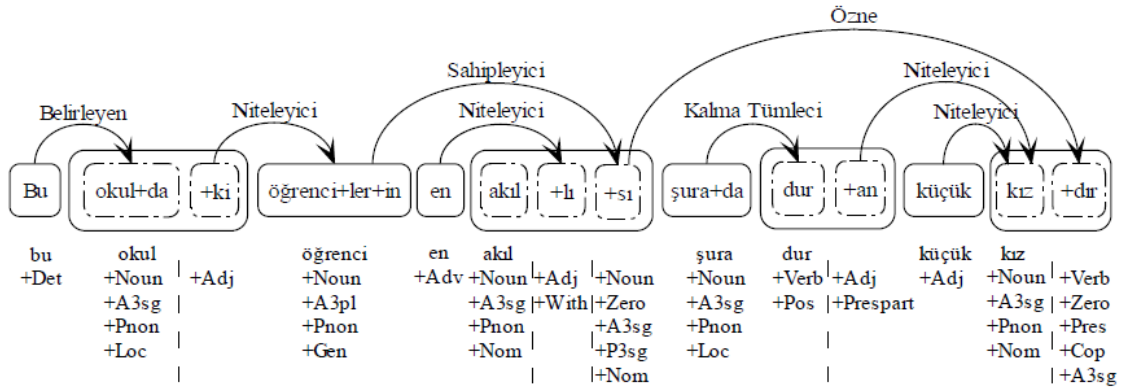
Birinci ÇK “oda” isimi ve bu isime ait biçimbilimsel özellikleri içermektedir. Bu özellikler, isimin tekil, iyelik eki almamış ve -de halinde olduğunu belirtmektedir. İkinci ÇK ise bu isimden türemiş “odada olmak” eylemini ve biçimbilimsel özelliklerini içermektedir. Eylem birinci tekil kişi eki almıştır ve şimdiki zamandır. Örnekte “küçük” olan, “odadayım” sözcüğü değil “oda” dır. “odadayım” isimden eyleme dönüşmüş bir sözcüktür. İki sözcük arasında kurulan bağlantı “odadayım” sözcüğünün eyleme dönüşmeden önceki isim halinden kaynaklanmaktadır. Bu durum sıfatların genel olarak isimlere bağlanması kuralından kaynaklanmaktadır. Bu cümlecikte, “Küçük olan nedir?” sorusunun cevabı bir sözcük sadece iye olarak bulunurlar. Bu ara ÇK'lerin arkalarından gelen ÇK'ye biçimbilimsel olarak bağlı oldukları varsayılır. Ancak bu bağılıklar özellikle belirtilmez. İye ÇK ise iye sözcüğün herhangi bir ÇK'si olabilir. Bir başka deyişle, bağılılık herhangi bir sözcüğün herhangi bir ÇK'sinde sonlanabilir [4].

Şekilden de görülebileceği gibi bir sözcükten sadece bir bağılılık oku çıkarken, birden fazla bağılılık oku girebilmektedir. Bir diğer deyişle, her sözcüğün sadece bir iyesi vardır ancak bir iye sözcüğün birden fazla uydusu olabilir. Birden çok ÇK içeren sözcüklere gelen bağılıklar sözcüğün farklı ÇK'lerinde sonlanabilir. Şekildeki “akıllısı” sözcüğü bu duruma güzel bir örnek teşkil etmektedir. “en” değil, bir çekim kümesidir [4].

Cümle içi ilişkilerin sözcüklerden daha küçük birimler arasında kurulabildiğine güzel bir örnek de Şekil 2.6’da gösterilmektedir. Şekil “Bu okuldaki öğrencilerin en akıllısı şurada duran küçük kızdır” cümlesi üzerinde çekim kümelerini ve bunlar arasında oluşan ilişkileri göstermektedir. Bağılıkların yönü uydu ÇK'den iye ÇK'ye doğru gösterilmiştir. Bağılılık türleri bağılılık oklarının üzerinde belirtilmektedir. Her

sözcüğe ait biçimbilimsel çözümleme ilgili sözcüğün altında Oflazer (1994)'in gösterimiyle verilmiştir. Bağlılıklar uydu sözcüğün sadece son ÇK'sinden çıkmaktadırlar. Bu nedenle, şekilde bağlılık çıkmayan bazı ÇK'ler bulunmaktadır (Örn. Okuldaki sözcüğünün ilk ÇK'si). Bu tip ÇK'ler bağlılıklarda sözcüğü “akıllısı” sözcüğünün ikinci ÇK'sine (“en akıllı”) bağlıdır. “Öğrencilerin” sözcüğü ise aynı sözcüğün üçüncü ÇK'sine bağlıdır (“öğrencilerin akıllısı”) [8].

Türkçe için bu tür bir yaklaşımın makine çevirisi, otomatik cümle analizi gibi üst düzey doğal dil işleme uygulamalarının başarımını önemli ölçüde arttırdığı çeşitli çalışmalarda gösterilmiştir [5] [9] [10].



Şekil 2.10. Türkçe cümlelerde cümle içi bağlılıklar

2.1.2. Eklerin Diziliş Kuralları

Kullanılacak herhangi bir yöntemden bağımsız olarak, Türkçe eklerin birbiri ardına geliş kuralları aşağıda iki farklı biçimde gösterilecektir. Bunlardan birincisi eklerin bir sözcük içerisinde soldan sağa doğru sözcük gövdesine hangi sırada ekleneceklerini sonlu durumlu makineler (SDM) ile gösterilmesidir. İkinci gösterim ise bu makinelerin ters çevrilmiş halleri yani sağdan sola halleridir. Birinci gösterim temel olarak Oflazer (1994)'in çalışmasındaki sonlu durum makineleri ile aynıdır. İkinci gösterim ise Cebiroğlu (2002)'nin çalışmasından alınmıştır. Cebiroğlu (2002)'nin geliştirmiş olduğu üstten-alta ek-atmalı ayrıştırma algoritmasında, bir sözcüğün ekleri sözcük sonundan başlanarak gövdeye doğru sözcükten birer birer çıkarılarak çözümleme yapılır. Bu çıkarma işlemi sonucunda elde edilen sözcük

gövde kabul edilebileceği gibi, bu varsayımı doğrulama için bir sözlük de kullanılabilir [8] [11].

Sadece bu makineleri kullanmak, Türkçe sözcüklerin tümünü doğru bir şekilde çözümleyecek yetkin bir çözümleyici oluşturmaya yeterli değildir. Yetkin bir çözümleyicinin çok fazla istisna durumu ve farklı kuralı ele alması ve Bölüm 4’de anlatılan kısıtlara uyması gerekecektir. Ancak bu SDM’ler ile hızlıca basit bir çözümleyici geliştirmek mümkündür [4].

Türkçede ekler sözcüklere bazı kurallara sadık kalınarak eklenirler. Benzer şekilde bir sözcük gövdesine birden fazla ek eklenmek istendiğinde bunların da belirli bir sıraya uyarak eklenmeleri gerekmektedir. Örneğin “kitap” sözcüğüne çoğul ve 1. Tekil kişi iyelik eki eklenmek istendiğinde, oluşacak sözcük “kitaplarım”dır. Bu durumda, “kitabımlar” şeklinde bir türeme yanlış kabul edilecektir. Buradan görülebilir ki, Türkçede iyelik ekleri genelde çoğul ekinden sonra gelmektedirler: arkadaşlarım, ödevleriniz vb. Ancak bazen kural dışı durumlar ortaya çıkabilir. Örn: halamlar, babaannemler vb [4].

Bölüm 1’de “kalem” örneği üzerinde gösterilen, bir sözcüğün birden fazla olası gövdesi olabilmesi durumunda ortaya çıkan belirsizlikler, sözcüğün üzerine farklı ekler eklendikçe ortadan kalkmaktadırlar. Bu fayda, Türkçede sadece bazı eklerin birbirlerinin ardına gelebilmesi sayesinde ortaya çıkmaktadır. Örneğin “kalem” sözcüğünün gövdesinin hem “kalem” hem de “kale” olabileceği önceki bölümlerde belirtilmiştir. Ancak “kalemlerim” sözcüğünün gövdesi sadece “kalem” olacaktır. “kale” gövdesi bu sözcük için mümkün değildir çünkü bir sözcük içerisinde hem iki tane iyelik eki olamaz hem de iyelik ekleri genelde çoğul ekinden sonra gelirler [4].

Bu bölümde, Türkçe dilbilgisinin ayrıntılarına girilmeyecektir. Bu tür bilgiler, herhangi bir Türkçe dil bilgisi kitabından kolayca elde edilebilirler. Ancak bu yapıların bilişimsel açıdan nasıl temsil edilebileceği ve bilgisayar tarafından nasıl işlenebileceği gibi konulara açıklık getirilmesi gerekmektedir. Bu amaçla, aşağıdaki bölümlerde Türkçe ekler gruplanarak, birbirleri ardına eklenişleri sonlu durum makineleri kullanılarak modelleneyecektir. Bu modelleme sayesinde, oluşturulan

makineler bir bilgisayar yazılımı sayesinde kolaylıkla işlenebilirler. Bu makineler tamamen doğru olmayıp, ekleniş kurallarının genel hallerini temsil etmektedirler. Bazı istisnai durumlar için makinelerde güncellemeler yapılması gerekebilir [4].

Takip eden bölümlerde, ekler aşağıdaki şekilde kümelerle ayrılarak incelenmektedir. Her bir bölümde, ilgili kümeye ait çekim eklerinin listesi bir tablo şeklinde verilecek, daha sonra sıra ile eklerin soldan sağa ve sağdan sola sıralanış kurallarını belirten sonlu durumlu makineler gösterilecektir [4].

2.1.3. İsim Çekim Ekleri Kümesi

Tablo 2.1’de ve 2.2’de isim soylu bir sözcüğe eklenebilecek çekim ekleri numaralandırılmış ve örneklenmiştir. Tablodaki “Ek” isimli sütunda yer alan ek tanımlamalarında kullanılan büyük harfler ve parantezler şu anlamlara gelmektedirler [4]:

A: a veya e harfi yerine

C: c veya ç harfi yerine

D: d veya t harfi yerine

H: ı, i, u veya ü harfi yerine

I: ı veya i harfi yerine

(): parantez içerisindeki harf ek içinde yer almayabilir.

Tablo 2.1. İsim çekim ekleri ek listesi

No	Ek	Açıklama	Örnek
1	-lAr	Çoğul	anne-ler
2	-(H)m	1. tekil kişi iyelik	anne-m
3	-(H)mHz	1. çoğul kişi iyelik	anne-miz
4	-(H)n	2. tekil kişi iyelik	anne-n
5	-(H)nHz	2. çoğul kişi iyelik	anne-niz
6	-(s)H	3. tekil kişi iyelik	anne-si
7	-lArI	1. çoğul kişi iyelik	anne-leri
8	-(y)H	-i hali	anne-yi

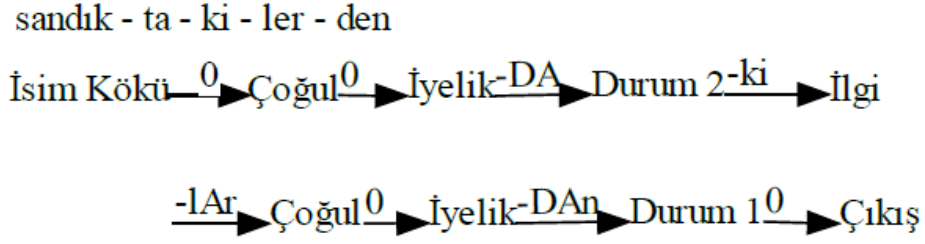
Örneğin, bu gösterimde “-lAr” şeklinde temsil edilen çoğul eki ekleneceği sözcükle birlikte uğrayacağı sesli uyumuna göre “-lar” veya “-ler” şeklinde görülebilir. Benzer şekilde “-(H)nIz” şeklinde gösterilen 2.çoğul kişi iyelik eki de ekleneceği sözcüğe göre “-nız, -niz, -nuz, -nüz, -ınız, -iniz, -unuz, -ünüz” (hırkanız, iğneniz, kokunuz, ütünüz, gözünüz vb.) şekillerinde görülebilir [4].

Şekil 2.7’deki makine, bir isim köküne getirilebilecek çekim eklerini ve bunların sıralanışındaki kuralları belirler. Bu SDM’de dikdörtgenler bir sözcüğün makine içerisinde bulunduğu durumları belirtirler. Daire içerisinde bulunan çıkışlar ise makinenin sonlanma durumlarını belirtirler. Makinenin bir tek başlangıç durumu “İsim Kökü” ve birden fazla sonlanma durumu vardır. Durumlar arası geçişler oklar ile belirtilmiştir. Oklar üzerinde yazılı ekler, makine içerisinde bir durumdan öteki duruma geçiş için sözcükte bulunması gereken ekleri belirtir. Örneğin “İsim Kökü” durumundan “İyelik (3)” durumuna geçilmesi için “-lArI” ekinin kök sonuna getirilmiş olması gerekir [4].

Aynı şekilde “İsim Kökü” durumundan “Çoğul” durumuna geçilmesi için şu ekler bulunmalıdır: “-lAr, 0”. Bu gösterimde, oklar üzerinde yazılı eklerin yanında bulunan ‘0’ bir ek değildir. Durumlar arası boş geçişleri gösterir. Böylece başlangıç durumu olan “İsim Kökü” durumundan “-(H)m, -(H)n, -(H)mHz, -(H)nHz” ekleri ile geçiş yapılabilir. Bu eklerden önce kökün sonuna “-lAr” çoğul eki getirilmesi zorunlu değildir [4].

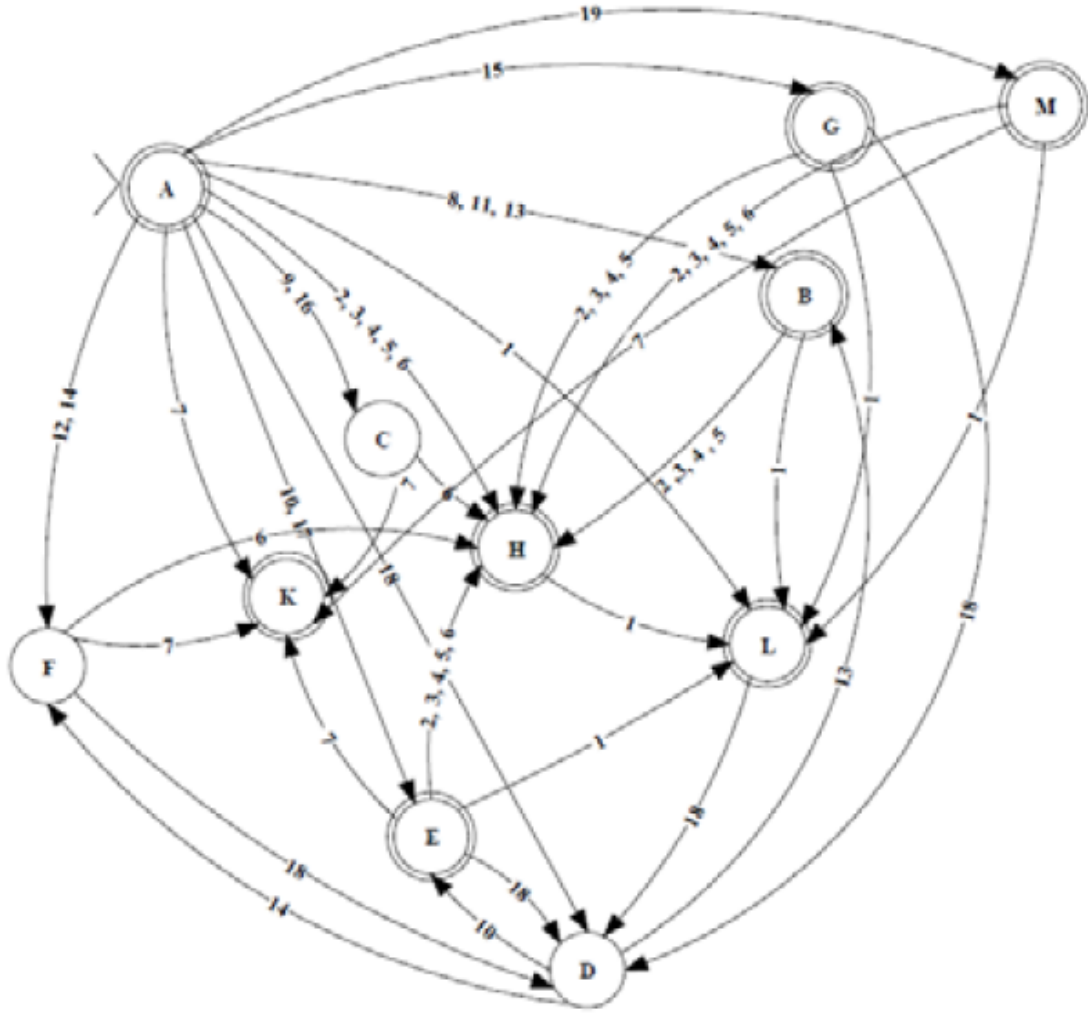
Bu sonlu durum makinesi ile yapılan analizde, isim soylu köke eklenen ekler ekleniş sıralarına göre incelenir ve “çıkış” durumuna ulaşılması ile sonlanır [4].

“sandıktakilerden” sözcüğünün İsim Çekim Ekleri Soldan Sağa SDM’si ile incelenmesi sırasında üzerinden geçilen durumlar aşağıda gösterilmiştir.



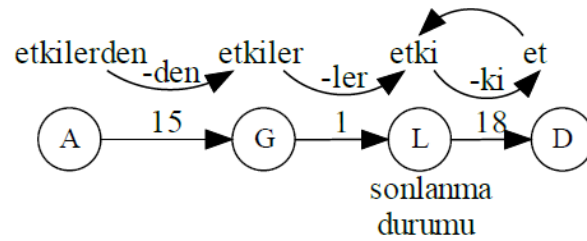
Şekil 2.12. "Sandıktakilerden" sözcüğünün soldan sağa incelemesi

Şekil 2.9’da Şekil 2.7’deki makinenin ters çevirme ve indirgeme işlemleri uygulanarak, gerekirci bir hale dönüştürülmesi ile ortaya çıkan ve eklerin sözcük sonundan başlanarak başa doğru sözcükten ayıklanmasını sağlayan sonlu durumlu makine gösterilmektedir. Bu gösterimde, durumlar daireler içerisinde, sonlanma durumu olabilen durumlar çift daire içerisinde, giriş durumları ise dairenin yanına büyük bir ok işareti konularak belirtilmiştir. Durumlar arası geçişlerin üzerinde yer alan sayılar, Tablo 2.1 İsim Çekim Ekleri tablosundaki eklerin numaralarıdır. “Sandıktakilerden” sözcüğünün bu makine ile incelenmesi sırasında üzerinden geçilecek durumlar aşağıdaki gibidir. İnceleme sonucu sandık isim kökü bulunmuştur [4].



Şekil 2.13. İsim çekim ekleri sağdan sola sonlu durum makinesi

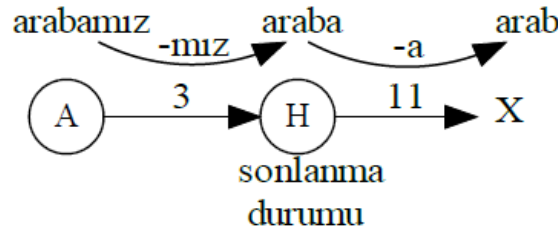
İnceleme sonuçlandığında üzerinde bulunan son durum bir sonlanma durumu olmalıdır. Aksi takdirde, üzerinden geçilmiş olan en son sonlanma durumu, incelemenin durdurulduğu yer olarak kabul edilir. Buna örnek olarak “etkilerden” sözcüğünün incelenmesi gösterilebilir [4].



Şekil 2.14. "Etkilerden" sözcüğünün sağdan sola incelenmesi

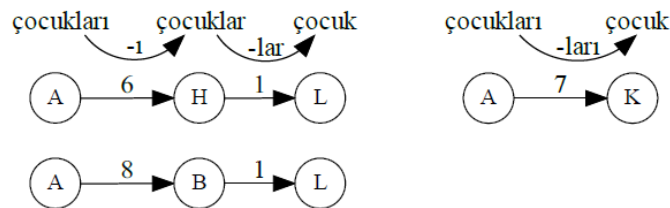
Bu örnekte en son durum olan “D” durumu bir sonlanma durumu değildir. Bu sebeple, kendisinden bir önceki sonlanma durumu olan “L”, incelemenin sonlandığı nokta olarak belirlenir. Bulunan isim kökü “et” değil, “etki” dir [4].

Sözcük üzerinde, sondan başa doğru inceleme yapılması sırasında, veri tabanında öncelikle olası bir ek bulunur. Üzerinde bulunulan durumdan, bu ek numarası ile başka bir duruma geçiş var ise bir sonraki duruma geçilir. Geçiş olmadığı durumda inceleme sonlanır. Arabamız sözcüğünün incelenmesi bu olaya bir örnek olarak gösterilebilir. Bu incelemede ulaşılan isim kökü “arab” değil “araba”dır [4].



Şekil 2.15. "Arabamız" sözcüğünün sağdan sola incelenmesi

Bazı sözcükler için birden çok biçimbilimsel analiz söz konusudur. Bu durum sözcük sonunda bulunan olası ekin birden fazla olduğu zamanlarda ortaya çıkar. Aşağıdaki örnekte *çocukları* sözcüğü için isim çekim ekleri kümesi için yaratılmış sonlu durum makinesi kullanılarak, üç farklı biçimbilimsel analiz yapılmıştır. Örnekte sözcük sonunda bulunan ilk ek ‘-ı’ veya ‘-ları’ eki olabilir. ‘-ı’ eki ‘6’ numaralı üçüncü tekil kişi iyelik eki olabileceği gibi, ‘8’ numaralı ismin –i hali eki de olabilir. Bu sebeple aynı sözcük için üç farklı inceleme yapılabilir [4].



Şekil 2.16. Birden çok biçimbilimsel analiz gerektiren sözcük örneği

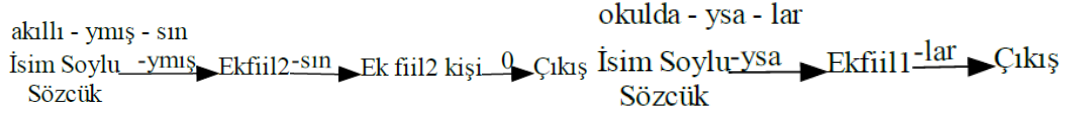
2.1.4. Ek-Eylem Ekleri Kümesi

Bu alt bölümde isim soylu sözcüklere eklenen ek-eylem ekleri için oluşturulan SDM tanıtılacaktır. Ek-eylem ekleri, bir isme getirilen zaman ve kişi eklerini içermektedirler. Tablo 2.3’de ek listesi açıklamalar ve örneklerle birlikte verilmektedir. İsim Çekim Ekleri bölümüne benzer şekilde bu bölümde de iki farklı gösterimde SDM verilecektir (Şekil 2.14 Ek-Eylem Ekleri Soldan Sağa Sonlu Durum Makinesi, Şekil 2.15 Ek-Eylem Ekleri Sağdan Sola Sonlu Durum Makinesi). Benzer şekilde, Şekil 7’deki SDM’de geçişlerde kullanılan sayılar, Tablo 2.3’deki ilk sütun olan ek numaralarını belirtmektedirler [4].

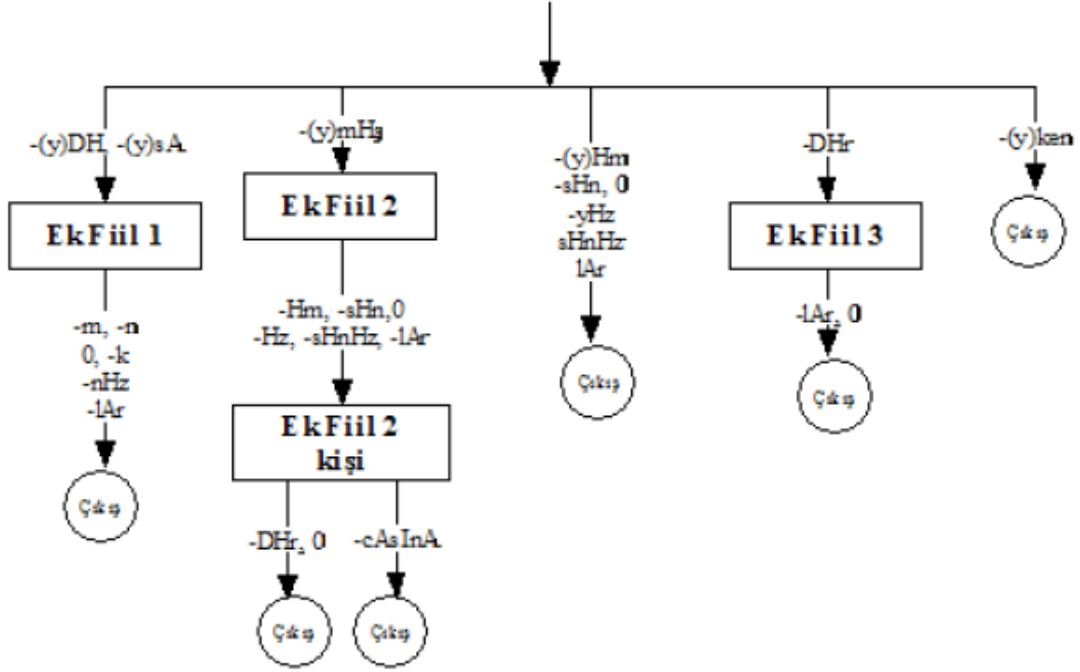
Tablo 2.3. Ek-eylem ekleri ek listesi

Ek No	Ek	Açıklama	Örnek
1	-(y)Hm	1. tekil kişi	akıllı-yım
2	-sHn	2. tekil kişi	akıllı-sın
3	-(y)Hz	1. çoğul kişi	akıllı-yız
4	-sHnHz	2. çoğul kişi	akıllı-sınız
5	-lAr	3. çoğul kişi	akıllı-lar
6	-m	1. tekil kişi ((y)DH ve (y)sA eklerinden sonra)	akıllı-ydı-m
7	-n	2. tekil kişi ((y)DH ve (y)sA eklerinden sonra)	akıllı-ysa-n
8	-k	1. çoğul kişi ((y)DH ve (y)sA eklerinden sonra)	akıllı-ysa-k
9	-nHz	2. çoğul kişi ((y)DH ve (y)sA eklerinden sonra)	akıllı-ydı-nız
10	-DHr	çevrik kip	akıllı-dır
11	-cAsInA	tarz zarfı	akıllı-ymış-casına
12	-(y)DH	di’li geçmiş zaman	akıllı-ydı
13	-(y)sA	dilek-şart kipi	akıllı-ysa
14	-(y)mHş	miş’li geçmiş zaman	akıllı-ymış
15	-(y)ken	zaman zarfı	akıllı-yken

Aşağıdaki örneklerde “akıllıymışsın” ve “okuldaysalar” sözcüklerinin soldan sağa SDM ile incelenmeleri gösterilmektedir [4].

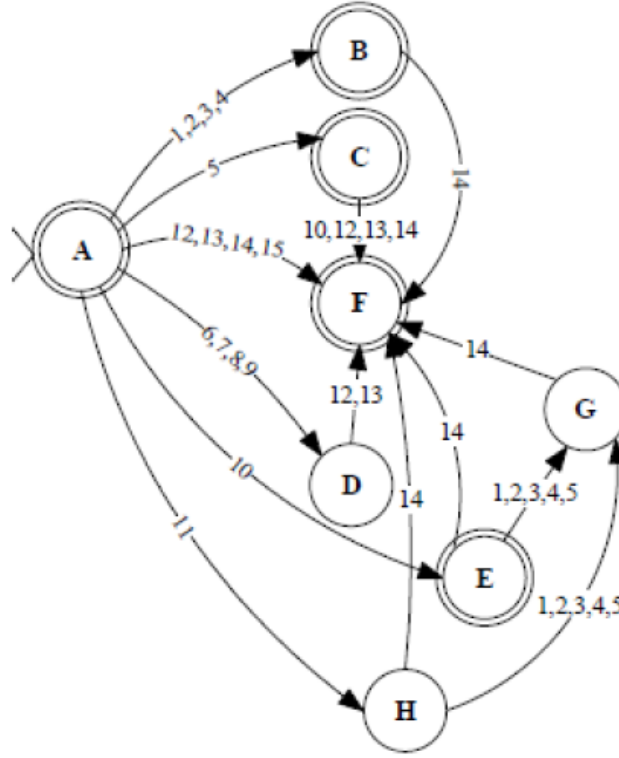


Şekil 2.17. “Akıllıymışsın” ve “okuldaysalar” sözcüklerinin soldan sağa SDM ile incelenmesi

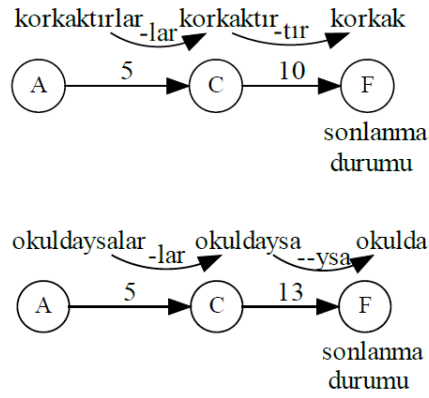


Şekil 2.18. Ek-eylem ekleri soldan sağa sonlu durum makinesi

Şekil 2.15’te gösterilen makine ile incelenen sözcüklerin sonunda ek-eylem eklerine rastlanması durumunda, bu ekler sözcükten çıkarılarak isim soylu sözcüğe ulaşılır. Elde edilen bu sözcük isim kökü olabileceği gibi, isim çekim ekleri almış isim soylu bir sözcük de olabilir. Aşağıdaki örneklerde incelenen “korkaktırlar” ve “okuldaysalar” sözcüklerinde elde edilen sonuç sırasıyla “korkak” ve “okulda” sözcükleridir. Bunlardan “korkak” bir isim kökü iken, okulda “okul” isim kökünün çekim eki almış halidir [4].



Şekil 2.19. Ek-eylem ekleri sağdan sola sonlu durum makinesi



Şekil 2.20. Sözcüklerin ek-eylem ekleri sağdan sola SDM ile incelenmesi

2.1.5. Eylem Zaman Ekleri Kümesi

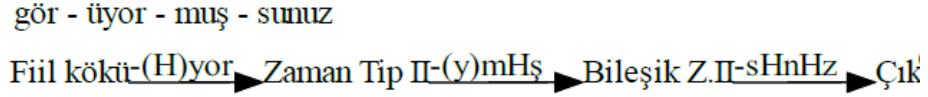
Eylem soylu sözcüklere eklenen çekim ekleri, eylem çekim ekleri ve eylem zaman ekleri olmak üzere iki kümeye ayrılabilir. Bu alt bölümde, eylem zaman ekleri kümesi için geliştirilen SDM tanıtılacaktır.

Üzerinde çalışılan bu kümede ekler, bir eyleme getirilebilecek zaman ve kişi eklerinden oluşmaktadır.

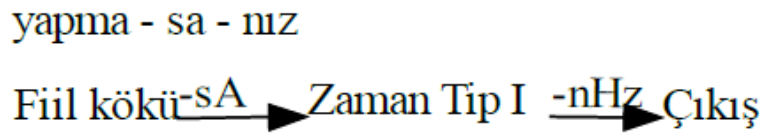
Tablo 2.4. Eylem zaman ekleri ek listesi

Ek No	Ek	Açıklama	Örnek
1	-(y)Hm	1. tekil kişi	gör-üyor-um
2	-sHn	2. tekil kişi	gör-üyor-sun
3	-(y)Hz	1. çoğul kişi	gör-üyor-uz
4	-sHnHz	2. çoğul kişi	gör-üyor-sunuz
5	-lAr	3. çoğul kişi	gör-üyor-lar
6	-mHş	miş'li geçmiş zaman	gör-müş
7	-(y)AcAk	gelecek zaman	gör-ecek
8	-(H)r	geniş zaman	gör-ür
9	-Ar	geniş zaman	iste-r
10	-(H)yör	şimdiki zaman	gör-üyor
11	-mAktA	sürerlilik	gör-mekte
12	-mAII	gereklilik	gör-meli
13	-m	1. tekil kişi	gör-dü-m
14	-n	2. tekil kişi	gör-dü-n
15	-k	1. çoğul kişi	gör-dü-k
16	-nHz	2. çoğul kişi	gör-dü-nüz
17	-DH	di'li geçmiş zaman	gör-dü
18	-sA	dilek-şart kipi	gör-se
19	-Ilm	1. çoğul kişi	gör-e-lim
20	-(y)A	istek kipi	gör-e
21	-(y)HnHz	2. çoğul kişi	gör-ünüz
22	-(y)Hn	2. tekil kişi	gör-ün
23	-sHnlAr	3. çoğul kişi	gör-sünler
24	-DHr	çevrik kip	gör-müş-sün-dür
25	-(y)DH	hikaye bileşik zaman	gör-müş-tü-m
26	-(y)sA	şart bileşik zaman	gör-müş-se-m
27	-(y)mHş	rivayet bileşik zaman	gör-meli-ymiş
28	-cAsInA	tarz zarfı	gör-müş-cesine
29	-(y)ken	zaman zarfı	gör-müş-ken

Şekil 2.20’de eylem zaman ekleri kümesi için oluşturulmuş soldan sağa SDM görülmektedir. Bu makinenin ‘olumsuz’ ve ‘fiil kökü’ durumları olmak üzere iki adet başlangıç durumu ve ‘çıkış’ olarak adlandırılan birden çok sonlanma durumu vardır. Aşağıda “görüyorsunuz” ve “yapmasanız” eylemlerinin bu makine ile incelenme evreleri gösterilmiştir [4].

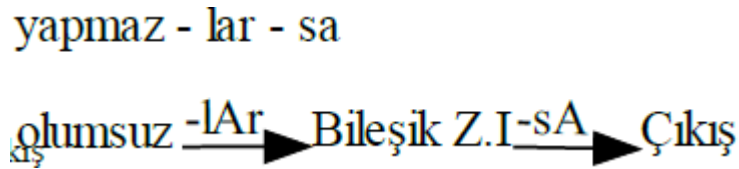


Şekil 2.21. "Görüyorsunuz" sözcüğünün incelenmesi

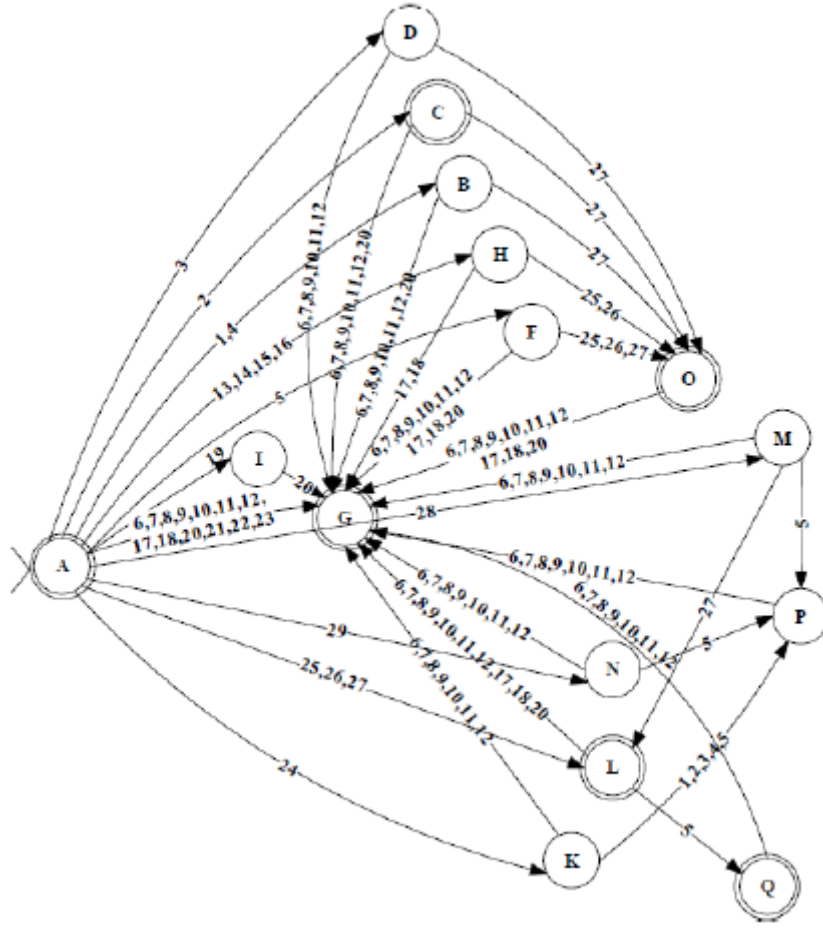


Şekil 2.22. "Yapmasanız" sözcüğünün incelenmesi

Makinede ‘olumsuz’ durumu olumsuz geniş zaman girişi için kullanılır. Olumlu durumlarda -(H)r veya -Ar ekleri ile oluşturulan geniş zaman, olumsuzluk halinde kendini -z eki ile gösterir. Aşağıdaki örnekte *yapmazlarsa* eylemi incelenmiştir. İnceleme sonucunda elde edilen ‘yapmaz’ sözcüğüdür. Bu sözcüğün üzerindeki -z geniş zaman olumsuzluk eki ve -mA olumsuzluk ekinin incelenmesi, bu SDM’nin görevi değildir [4].

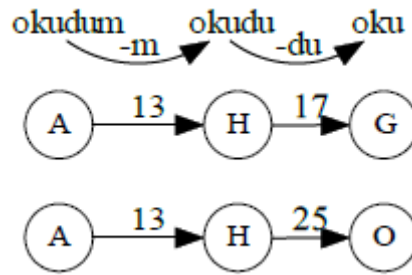


Şekil 2.23. "Yapmazlarsa" sözcüğünün incelenmesi



Şekil 2.25. Eylem zaman ekleri sağdan sola sonlu durum makinesi

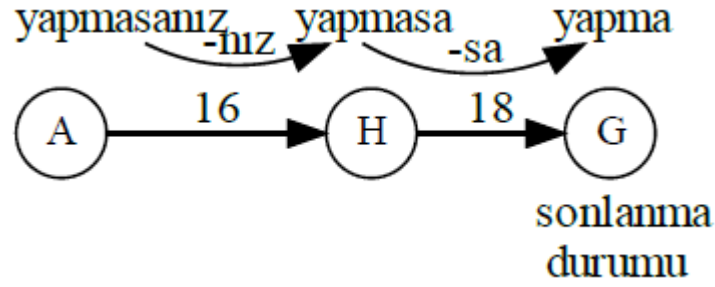
Aşağıda “okudum”, “yapmanız”, “yapmazlarsa” ve “gezermişiz” eylemleri için inceleme aşamaları görülmektedir [4].



Şekil 2.26. "Okudum" sözcüğünün incelemesi

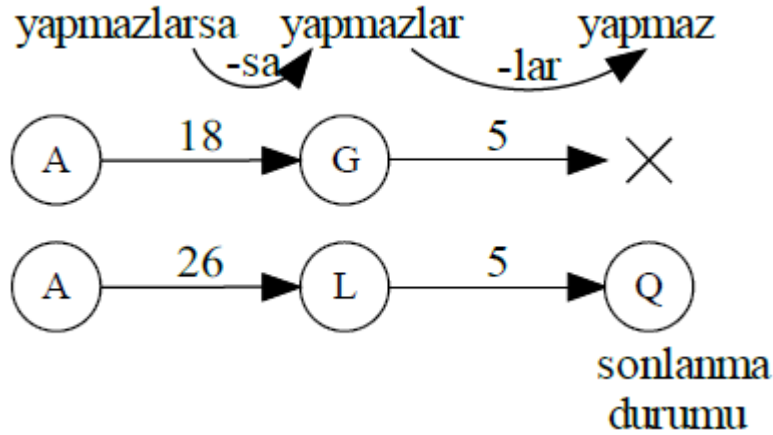
Bir eyleme bileşik zaman ekinin getirilebilmesi için kendisinden önce gelen bir normal zaman ekinin bulunması şarttır. Yukarıdaki örnekte –du eki veri tabanında arandığında iki farklı sonuç elde edilir: –di’li geçmiş zaman eki veya hikaye bileşik

zaman eki. Bu sonuçların her ikisi de sonlanma durumuna geçiş sağlar. Ancak, ‘O’ durumu olumsuz eylemler için bir sonlanma durumu olduğundan, ilk inceleme doğru olarak kabul edilmesi gerektir. SDM eylemin olumsuz olup olmadığı kontrolünü yapamaz. Bu görev bir sonraki SDM olan Eylem Çekim Ekleri Kümesi’ne aittir. İki farklı inceleme sonucu üretilir. Diğer SDM’ler tarafından yapılan incelemelerde eylemin olumsuz olduğu belirlenemez ise ‘O’ çıkışlı inceleme kaydı silinir [4].



Şekil 2.27. "Yapmasanız" sözcüğünün incelemesi

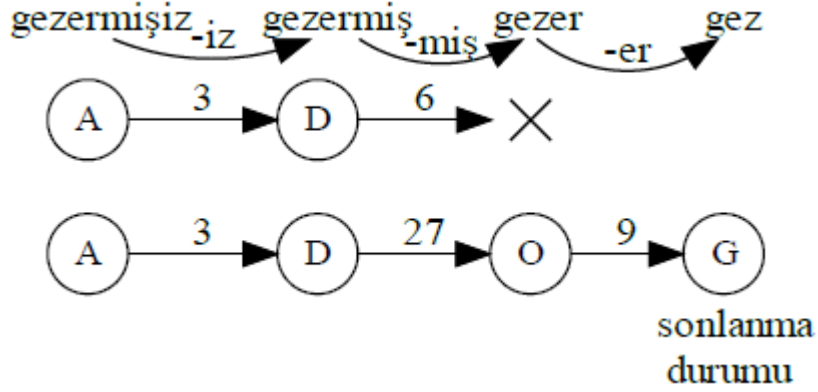
Yukarıdaki örnekte olumsuz bir eylem için inceleme görülmektedir. Kullanılan –sa eki hem dilek-şart eki hem de şart bileşik zaman eki olabilir. Bu yüzden veri tabanında arama yapıldığında birden fazla sonuç geriye döner: ‘18’ ve ‘26’.



Şekil 2.28. "Yapmazlarsa" sözcüğünün incelemesi

Yukarıdaki örnekte birinci incelemede analiz ‘G’ durumuna kadar ilerleyebilir, ‘5’ numaralı ek ile bu durumdan başka herhangi bir duruma geçiş yoktur. Elde edilen

sözcük “yapmazlar”dır. İkinci inceleme de ise ‘Q’ sonlanma durumuna kadar ulaşılmıştır, elde edilen sözcük “yapmaz”dır [4].



Şekil 2.29. "Gezermişiz" sözcüğünün incelemesi

Yukarıdaki örnekte, birinci inceleme ‘D’ durumuna kadar ilerleyebilir, bu durum bir sonlanma durumu olmadığından kendisinden önceki ilk sonlanma durumu, son durum olarak kabul edilir. Sözcük birinci incelemeden herhangi bir değişikliğe uğramadan çıkar. İkinci inceleme de ise ‘G’ sonlanma durumuna ulaşılır, elde edilen sözcük gez eylem köküdür [4].

2.1.6. Eylem Çekim Ekleri Kümesi

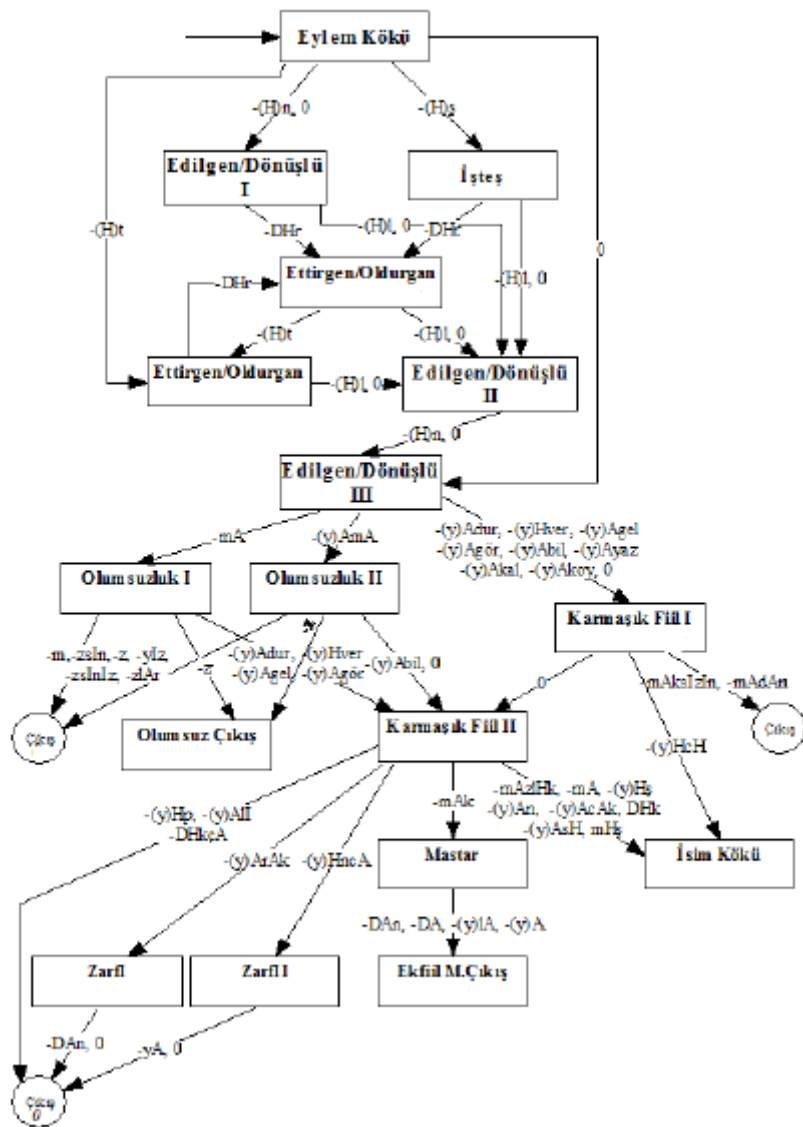
Kişi ve zaman ekleri dışında, eylemlere eklenen çekim ekleri, eylem çekim ekleri kümesi altında toplanmıştır. Bu küme, karmaşık fiil eklerini, tasvir fiil eklerinin ve çatı eklerini içerir. Bu alt bölümde eylem çekim ekleri kümesi için oluşturulmuş SDM tanıtılacaktır. Şekil 10’da verilen eylem çekim ekleri kümesi için oluşturulmuş soldan sağa SDM’de “Eylem Kökü” durumu makinenin giriş durumudur. Eylem çekim ekleri SDM’sinde, sonlanma durumları sadece “çıkış” durumları değildir. “Çıkış” durumlarının dışında, kendisinden farklı SDM’lere geçişi sağlayan sonlanma durumları vardır. Bunlar “Karmaşık Fiil II” , “İsim Kökü”, “Mastar”, “Ekfiil M. Çıkış” ve “Olumsuz Çıkış” durumlarıdır. Bunların her biri SDM’lere geçişi sağlar. Farklı çıkışları tanıtmak üzere aşağıda “görüşmezler”, “gönderilebilir”, “okuyucular”, “okumak”, “okumaktadır” ve “okuyamazmışım” sözcükleri üzerinde incelemeler yapılmıştır [4].

Tablo 2.5. Eylem çekim ekleri ek listesi

Ek No	Ek	Açıklama	Örnek
1	–m	1. tekil kişi	yap-ma-m
2	–zsIn	2. tekil kişi	yap-ma-zsın
3	–z	3. tekil kişi	yap-ma-z
4	–yIz	1. çoğul kişi	yap-ma-yız
5	–zsInIz	2. çoğul kişi	yap-ma-zsınız
6	–zIAr	3. çoğul kişi	yap-ma-zlar
7	–mA	olumsuzluk	yap-ma
8	–(y)AmA	olumsuzluk	yap-ama
9	–(y)Adur	sürerlik bileşik eylem kipi	yap-adur
10	–(y)Hver	tezlik bileşik eylem kipi	yap-ıver
11	–(y)Agel	sürerlik bileşik eylem kipi	ol-agel-di
12	–(y)Agör	sürerlik bileşik eylem kipi	iste-yegör-sün
13	–(y)Abil	yeterlik bileşik eylem kipi	yap-abil
14	–(y)Ayaz	yaklaşma bileşik eylem kipi	düş-eyaz-dı-m
15	–(y)Akal	sürerlik bileşik eylem kipi	don-akal
16	–(y)Akoy	sürerlik bileşik eylem kipi	alı-koy-sun
17	–mAk	mastar	yap-mak
18	–(y)HcH	görev eki	yap-ıcı
19	–(y)Hp	zarf eki	yap-ıp
20	–(y)AlI	zarf eki	yap-alı
21	–DHkçA	zarf eki	yap-tıkça
22	–(y)ArAk	zarf eki	yap-arak
23	–(y)HncA	zarf eki	yap-ınca
24	–DAn	zarf eki	yap-arak-tan
25	–yA	zarf eki	yap-ınca-ya
26	–(y)An	sıfat fiil	oku-yan
27	–(y)AcAk	fiilden isim yapma eki	oku-yacak
28	–(y)AsI	sıfat fiil	oku-yası
29	–DHk	sıfat fiil	oku-duk
30	–mHş	sıfat fiil	oku-muş

Tablo 2.6. Eylem çekim ekleri ek listesi

31	–mAzllk	fiilden isim yapma eki	oku-ma-mazlık
32	–mA	mastar eki	oku-ma
33	–(y)Hş	fiilden isim yapma eki	oku-yuş
34	–DAn	-den hali	oku-mak-tan
35	–DA	-de hali	oku-mak-ta
36	–(y)lA	birliktelik	oku-ma-yla
37	–(y)A	-e hali	oku-ma-ya
38	–mAksızln	zarf eki	oku-maksızın
39	–mAdAn	zarf eki	oku-madan
40	–(H)n	edilgen dönüşlü	oku-n-mak
41	–(H)ş	işteş	gör-üş-mek
42	–(H)l	edilgen dönüşlü	gör-ül-mek
43	–DHr	oldurgan	oku-t-tur-mak
44	–(H)t	ettirgen	oku-t-mak



Şekil 2.30. Eylem çekim ekleri soldan sağa sonlu durum makinesi

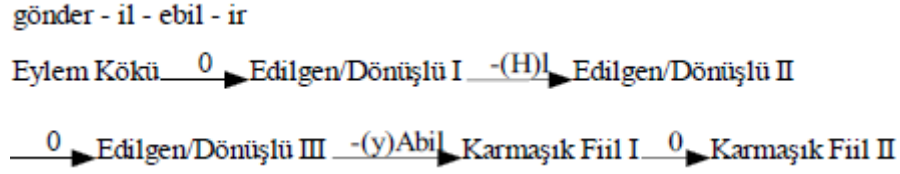
gör - üş - me - zler

Eylem Kökü $\xrightarrow{-(H)\text{ş}}$ İsteş $\xrightarrow{0}$ Edilgen/Dönüşlü II $\xrightarrow{0}$ Edilgen/Dönüşlü III

$\xrightarrow{-mA}$ Olumsuzluk I $\xrightarrow{-zlAr}$ Çıkış

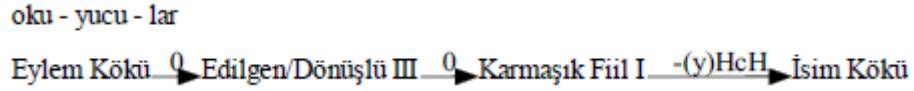
Şekil 2.31. "Görüşmezler" sözcüğünün incelemesi

“Çıkış” normal sonlanma durumudur, inceleme bu SDM’de başlar ve sona erir [4].



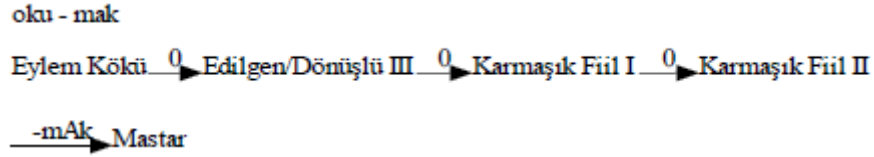
Şekil 2.32. "Gönderilebilir" sözcüğünün incelenmesi

“Karmaşık Fiil II” durumu, bu SDM’den eylem zaman ekleri SDM’sine geçişi sağlayan kapıdır. Yukarıdaki örnekte “-ir” eki, eylem çekim eklerinde incelenecek bir ek değildir. “Karmaşık Fiil II” sonlanma durumuna ulaşan inceleme, buradan eylem zaman ekleri SDM’sine geçer ve devam eder [4].



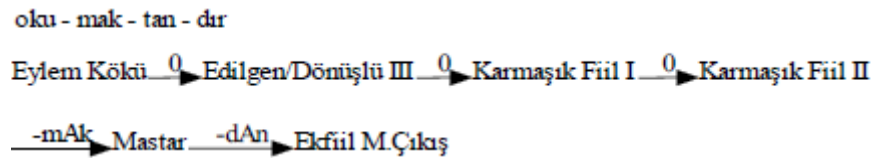
Şekil 2.33. "Okuyucular" sözcüğünün incelenmesi

“İsim Kökü” sonlanma durumu bu SDM’sinden isim çekim ekleri SDM’sine geçişi sağlar. İnceleme bu aşamadan sonra isim çekim ekleri SDM’sinden devam eder [4].



Şekil 2.34. "Okumak" sözcüğünün incelenmesi

“Master” normal sonlanma durumudur, inceleme bu SDM’de başlar ve sona erer [4].



Şekil 2.35. "Okumaktadır" sözcüğünün incelenmesi

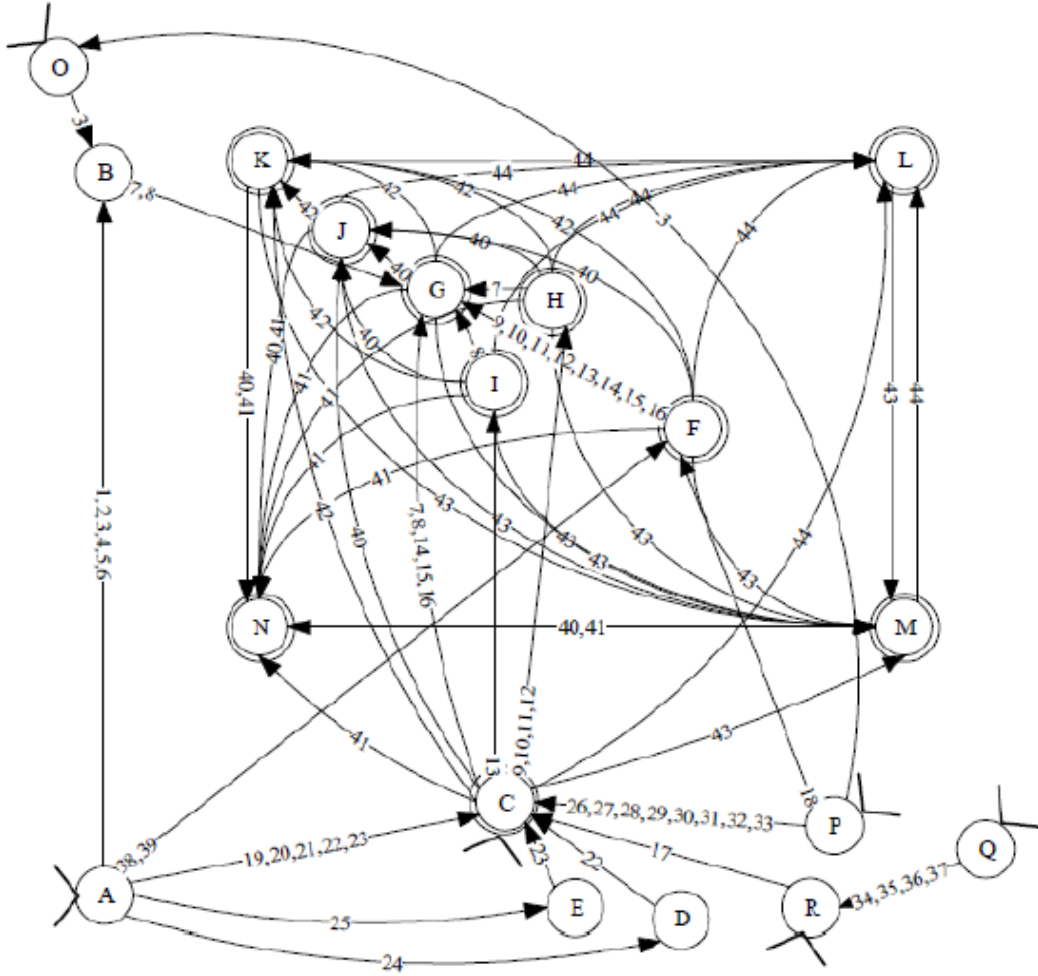
“Ekfiil M.Çıkış” sonlanma durumu bu SDM’den ekeylem ekleri SDM’sine geçişi sağlar. Bu duruma gelmiş inceleme ekeylem ekleri SDM’sinden devam eder. “-dır” eki ekeylem ekleri SDM’sinde incelenir [4].

oku - yama - z - mış - ım

Eylem Kökü 0 Edilgen/DönüşlüIII $-(y)AmA$ OlumsuzlukII -2 Olumsuz Çıkış

Şekil 2.36. "Okuyamazmıım" sözcüğünün incelemesi

“Olumsuz Çıkış” durumu bu SDM’den, eylem zaman ekleri SDM’sine geçişi sağlar. Örnekteki “-mış” ve “-ım” ekleri eylem zaman ekleri SDM’si tarafından incelenecektir [4].



Yukarıda örnekler ile açıklanan, eylem çekim ekleri SDM'sinden diğer SDM'lere geçişler şöyledir:

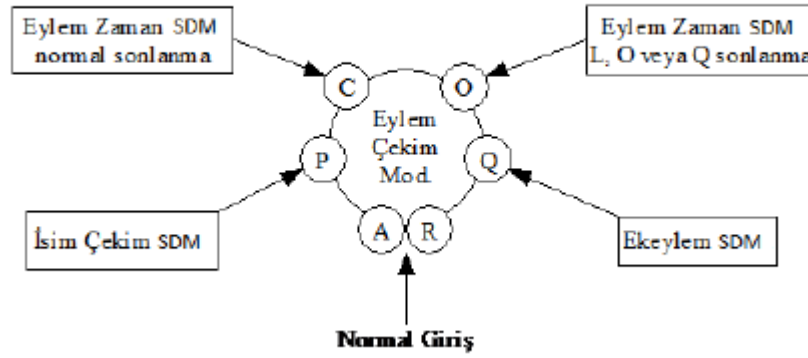
“Olumsuzluk Çıkış” durumu → Eylem zaman ekleri SDM'si

“Karmaşık Fiil I” durumu → Eylem zaman ekleri SDM'si

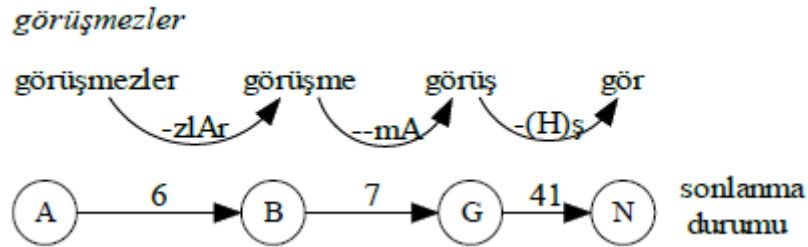
“İsim Kökü” durumu → İsim çekim ekleri SDM'si

“Ekfiil M.Çıkış” durumu → Ekeylem ekleri SDM'si

Eylem çekim ekleri için oluşturulmuş sağdan sola SDM Şekil 2.33'de gösterilmektedir. Bu makineden önce birçok farklı SDM çalışabileceğinden, makineye girişler öncesinde çalışan SDM'ye göre farklılık göstermektedirler. Bu girişlerin neler olduğu Şekil 2.34'de gösterilmektedir [4].

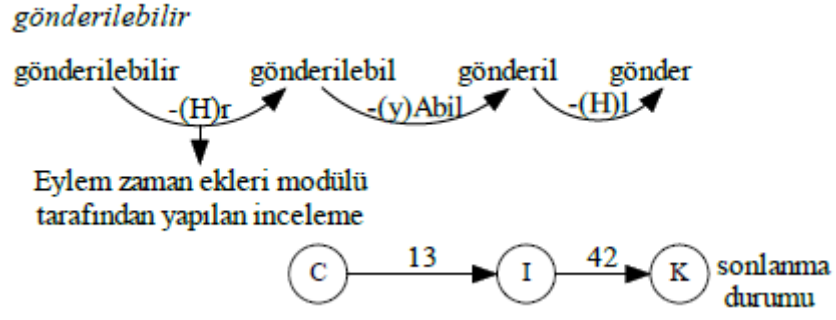


Şekil 2.38. Sonlu durum makinesi girişleri



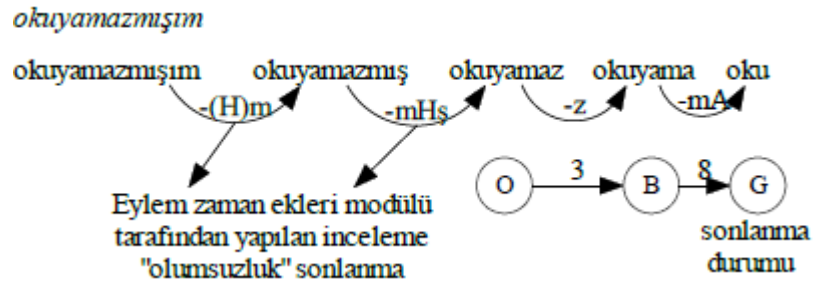
Şekil 2.39. "Görüşmezler" sözcüğünün incelemesi

Yukarıdaki örnekte eylem çekim ekleri SDM'sinden önce inceleme yapan başka bir SDM olmadığından giriş 'A' durumundandır [4].



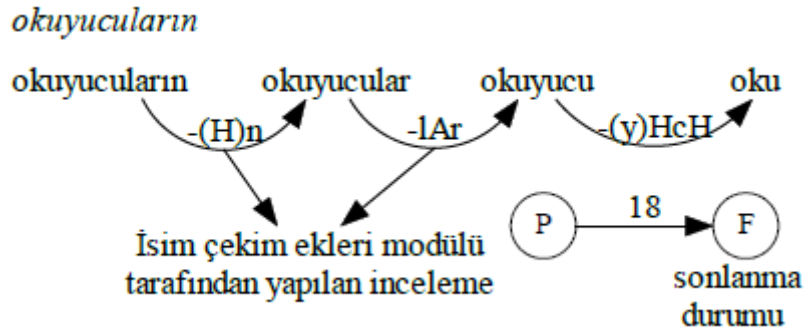
Şekil 2.40. "Gönderebilir" sözcüğünün incelenmesi

Yukarıdaki örnekte, inceleme yapan bir önceki SDM eylem zaman ekleri SDM'si olduğundan, başlangıç olarak 'C' durumu kabul edilir. Bu SDM'deki inceleme "gönderebil" sözcüğünden başlar [4].



Şekil 2.41. "Okuyamazmışım" sözcüğünün incelenmesi

Yukarıdaki örnekte, inceleme yapan bir önceki SDM, eylem zaman ekleri SDM'si ve sonlanma durumu olumsuzluk olduğundan, başlangıç durumu olarak 'O' kabul edilir. Bu SDM'deki inceleme "okuyamaz" sözcüğünden başlar [4].



Şekil 2.42. "Okuyucuların" sözcüğünün incelenmesi

numaralandırılması yeterli görülmüştür. Tablo 2.7’de bazı yapım ekleri verilmiştir [11].

Tablo 2.7. Yapım ekleri ek listesi

Ek No	Ek	Açıklama	Örnek
1	–lHk	isimden isim türeten	yol-luk
2	–CH	isimden isim türeten	yol-cu
3	–CHk	isimden isim türeten	kabar-cık
4	–lAş	isimden fiil türeten	veda-laş-mak
5	–lA	isimden isim türeten	uğur-la-mak
6	–lAn	isimden isim türeten	uğur-lan-mak
7	–CA	isimden sıfat türeten	insan-ca
8	–lH	isimden sıfat türeten	araba-lı
9	–sHz	isimden sıfat türeten	araba-sız

BÖLÜM 3. SONUÇLAR VE ÖNERİLER

Bitirme çalışması boyunca Türkçe'nin biçimbilimsel analizi yapılmış, Türkçe'deki eklerin yapısı incelenmiştir. Türkçe gibi sondan eklemeli dillerde kök bulunması hususu, eklerin diziliş kurallarının belirlenmesi zorunluluğunu beraberinde getirmiştir. Derleyici aşamalarında yer alan sözcüksel, sözdizimsel (sentaktik) ve anlamsal (semantik) analiz tamamlanmış, eklerin diziliş kurallarına göre sonlu durum makineleri tasarlanmıştır.

Bitirme çalışmasında gerçekleştirilen analizlerden yola çıkarak ilerleyen aşamalarda geliştirilebilecek bir program tasarlanmış ve gerçekleştirilmiştir. Türkçe bir sözcüğün kökünü tespit eden ve eklerini belirleyen program, Türkçe'nin bilgisayar tarafından anlaşılabilirliğini arttırmıştır. Geliştirilen programın, ilerleyen aşamalarda çeviri kanallarına, arama motorlarına, sözcük işlem programlarına ve yazılım geliştirme ortamlarına uyarlanması hedeflenebilir.

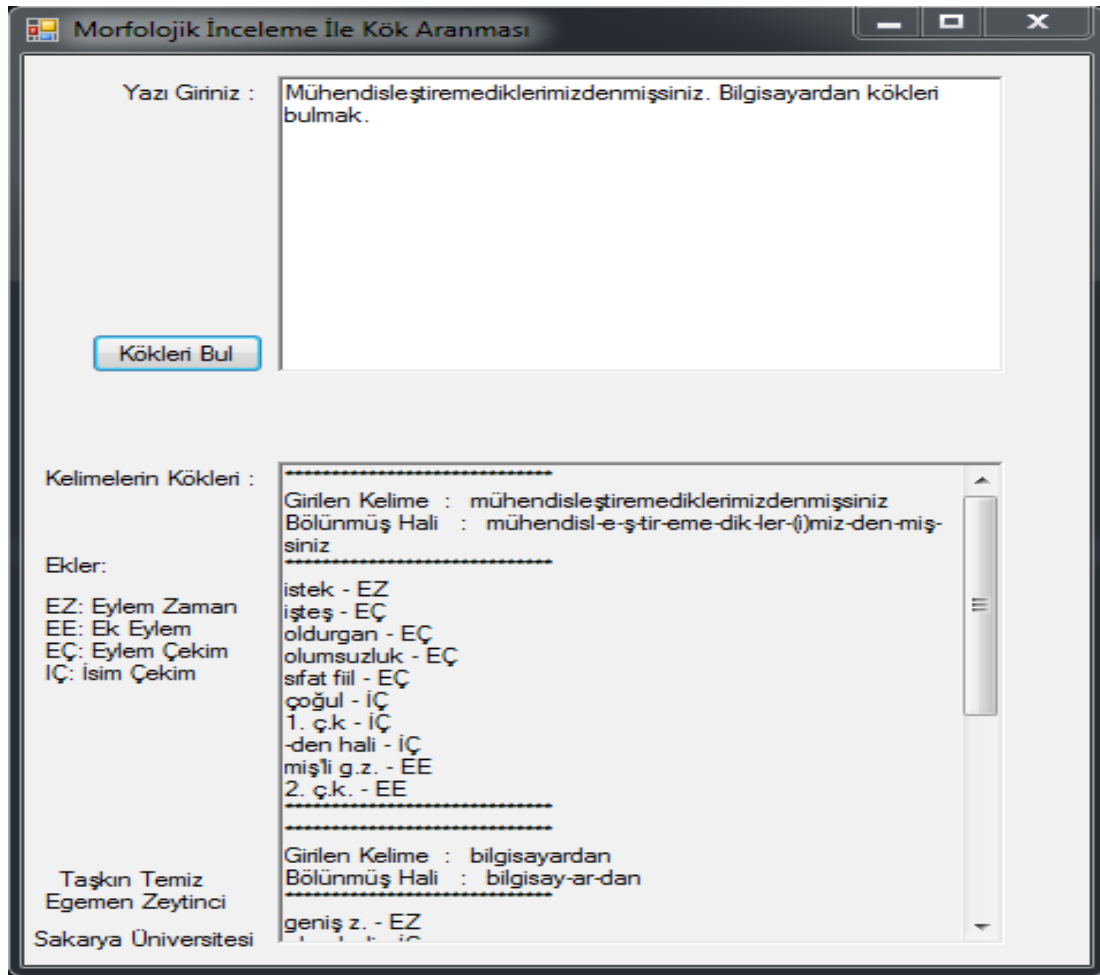
KAYNAKLAR

- [0] ZENGİN, E., Derleyicilerin Yapısı, *Bilişim Dergi*, Sayı: 14, Mayıs, 2010.
- [0] <http://e-bergi.com/y/Derleyiciler>, Erişim Tarihi: 16.05.2016.
- [0] <http://www.bilisimdergi.com/Derleyicilerin-Yapisi---2-15-10.html>, Erişim Tarihi: 16.05.2016.
- [0] ERYİĞİT, G. Biçimbilimsel Çözümleme. *Türkiye Bilişim Vakfı Bilgisayar Bilimleri ve Mühendisliği Dergisi*, 6.2, 33-55, 2013.
- [0] OFLAZER, K., TÜR, G. Statistical Morphological Disambiguation for Agglutinative Languages. *Journal of Computers and Humanities*, 36.4, 2002.
- [0] OFLAZER, K. Dependency parsing with an extended finite-state approach. *Computational Linguistic* , 29.4, 515-544, 2003.
- [0] ERYİĞİT, G. Statistical Dependency Parsing of Turkish. *EACL*. Trento, 2006.
- [0] OFLAZER, K. Two-level description of Turkish morphology. *Literary and Linguistic Computing* , 9.2, 1994.
- [9] ERYİĞİT, G., Nivre, J., Oflazer, K. Dependency Parsing of Turkish. *Computational Linguistic* , 34.3, 357–389, 2008.
- [10] EL-KAHLOUT, İ. D., OFLAZER, K. Exploiting Morphology and Local Word Reordering in English to Turkish Phrase-based Statistical. *IEEE Transactions on Audio, Speech and Language Processing*, 2009.
- [11] CEBİROĞLU, G. Sözlüksüz Köke Ulaşma Yöntemi, *Yüksek Lisans Tezi*. İstanbul Teknik Üniversitesi, (2002).

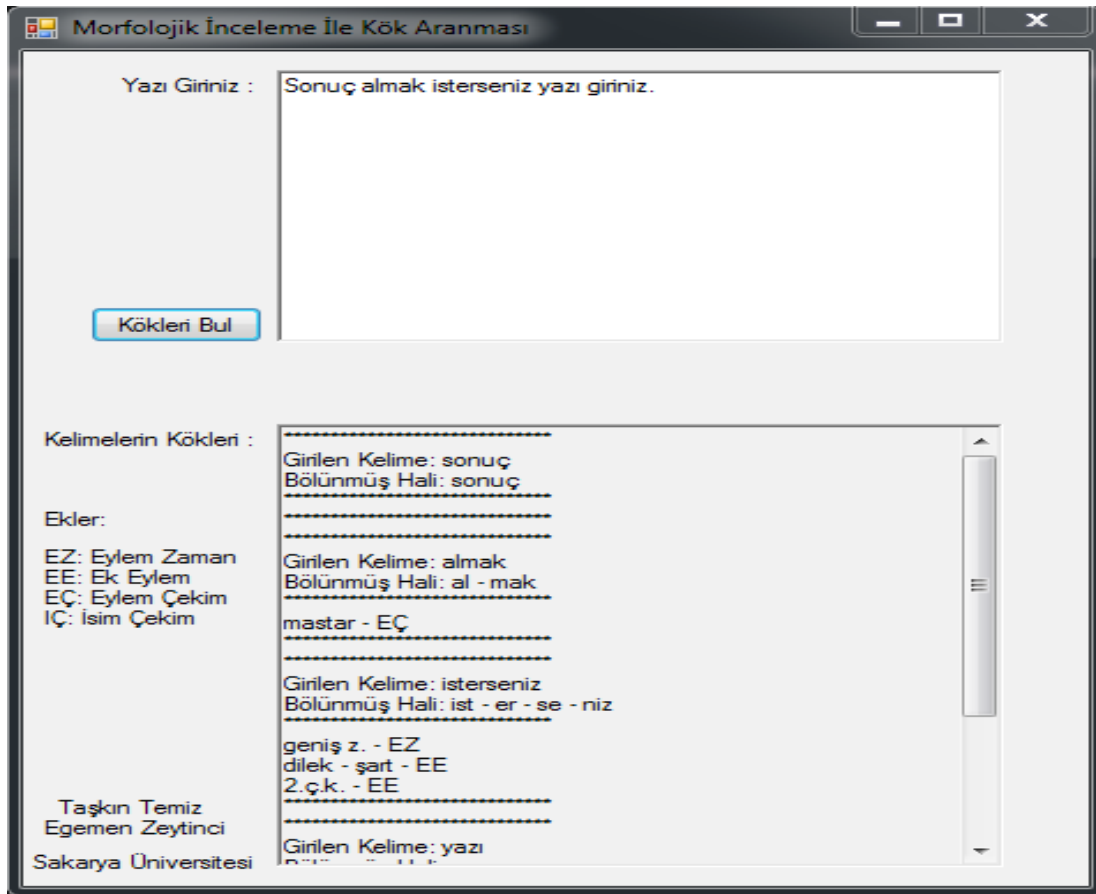
EKLER

EK A: Aşağıdaki program çıktısında

“Mühendisleştiremediklerimizdenmişsiniz.” sözcüğünün ve
 “Bilgisayardan kökleri bulmak.” cümlesinin biçimbirimsel analizi
 yapılmıştır ve programın bulduğu ekler isimleri ile birlikte ekrana
 yazdırılmıştır.



EK B: Aşağıdaki program çıktısında “Sonuç almak isterseniz yazı giriniz.” cümlesinin biçimbirimsel analizi yapılmıştır ve programın bulunduğu ekler isimleri ile birlikte ekrana yazdırılmıştır.



ÖZGEÇMİŞ

Taşkın Temiz, 10.05.1992 de İstanbul'da doğdu. İlk ve orta eğitimini Küçükçekmece'de tamamladı. Lise eğitimini öncelikle Hadımköy Örfi Çetinkaya Anadolu Lisesi'nde sonrasında Bakırköy Yahya Kemal Beyatlı Lisesi'nde alarak 2010 yılında tamamladı. 2010 yılında başladığı Sakarya Üniversitesi Bilgisayar Mühendisliği Bölümü'de eğitimini sürdürmekte.

Egemen Zeytinci, 09.11.1993 de Aydın'da doğdu. İlk ve orta eğitimini Aydın'da tamamladı. Lise eğitimini öncelikle Aydın Emel Mustafa Uşaklı Anadolu Lisesi'nde sonrasında Aydın Atatürk Anadolu Lisesi'nde alarak 2011 yılında tamamladı. 2012 yılında başladığı Sakarya Üniversitesi Bilgisayar Mühendisliği Bölümü'de eğitimini sürdürmekte.

BSM 498 BİTİRME ÇALIŞMASI DEĞERLENDİRME VE SÖZLÜ SINAV TUTANAĞI

KONU : TÜRKÇE DİLİNİN MORFOLOJİK İNCELEMESİ VE BİÇİMSEL OLARAK TANIMLANMASI

ÖĞRENCİLER (Öğrenci No/AD/SOYAD): G101210009/TAŞKIN/TEMİZ
B121210057/EGEMEN/ZEYTİNCİ

Değerlendirme Konusu	İstenenler	Not Aralığı	Not
Yazılı Çalışma			
Çalışma klavuza uygun olarak hazırlanmış mı?	x	0-5	
Teknik Yönden			
Problemin tanımı yapılmış mı?	x	0-5	
Geliştirilecek yazılımın/donanımın mimarisini içeren blok şeması (yazılımlar için veri akış şeması (dfd) da olabilir) çizilerek açıklanmış mı?			
Blok şemadaki birimler arasındaki bilgi akışına ait model/gösterim var mı?			
Yazılımın gereksinim listesi oluşturulmuş mu?			
Kullanılan/kullanılması düşünülen araçlar/teknolojiler anlatılmış mı?			
Donanımların programlanması/konfigürasyonu için yazılım gereksinimleri belirtilmiş mi?			
UML ile modelleme yapılmış mı?			
Veritabanları kullanılmış ise kavramsal model çıkarılmış mı? (Varlık ilişki modeli, noSQL kavramsal modelleri v.b.)			
Projeye yönelik iş-zaman çizelgesi çıkarılarak maliyet analizi yapılmış mı?			
Donanım bileşenlerinin maliyet analizi (prototip-adetli seri üretim vb.) çıkarılmış mı?			
Donanım için gerekli enerji analizi (minimum-uyku-aktif-maksimum) yapılmış mı?			
Grup çalışmalarında grup üyelerinin görev tanımları verilmiş mi (iş-zaman çizelgesinde belirtilebilir)?			
Sürüm denetim sistemi (Version Control System; Git, Subversion v.s.) kullanılmış mı?			
Sistemin genel testi için uygulanan metotlar ve iyileştirme süreçlerinin dökümü verilmiş mi?			
Yazılımın sızma testi yapılmış mı?			
Performans testi yapılmış mı?			
Tasarımın uygulamasında ortaya çıkan uyumsuzluklar ve aksaklıklar belirtilerek çözüm yöntemleri tartışılmış mı?			
Yapılan işlerin zorluk derecesi?	x	0-25	
Sözlü Sınav			
Yapılan sunum başarılı mı?	x	0-5	
Soruları yanıtlama yetkinliği?	x	0-20	
Devam Durumu			
Öğrenci dönem içerisindeki raporlarını düzenli olarak hazırladı mı?	x	0-5	
Diğer Maddeler			
Toplam			

DANIŞMAN (JÜRİ ADINA): PROF. DR. NEJAT YUMUŞAK

DANIŞMAN İMZASI:

BSM 498 BİTİRME ÇALIŞMASI
DEĞERLENDİRME VE SÖZLÜ SINAV TUTANAĞI

KONU :

ÖĞRENCİLER (Öğrenci No/AD/SOYAD):

Değerlendirme Konusu	İstenenler	Not Aralığı	Not
Yazılı Çalışma			
Çalışma klavuza uygun olarak hazırlanmış mı?	x	0-5	
Teknik Yönden			
Problemin tanımı yapılmış mı?	x	0-5	
Geliştirilecek yazılımın/donanımın mimarisini içeren blok şeması (yazılımlar için veri akış şeması (dfd) da olabilir) çizilerek açıklanmış mı?			
Blok şemadaki birimler arasındaki bilgi akışına ait model/gösterim var mı?			
Yazılımın gereksinim listesi oluşturulmuş mu?			
Kullanılan/kullanılması düşünülen araçlar/teknolojiler anlatılmış mı?			
Donanımların programlanması/konfigürasyonu için yazılım gereksinimleri belirtilmiş mi?			
UML ile modelleme yapılmış mı?			
Veritabanları kullanılmış ise kavramsal model çıkarılmış mı? (Varlık ilişki modeli, noSQL kavramsal modelleri v.b.)			
Projeye yönelik iş-zaman çizelgesi çıkarılarak maliyet analizi yapılmış mı?			
Donanım bileşenlerinin maliyet analizi (prototip-adetli seri üretim vb.) çıkarılmış mı?			
Donanım için gerekli enerji analizi (minimum-uyku-aktif-maksimum) yapılmış mı?			
Grup çalışmalarında grup üyelerinin görev tanımları verilmiş mi (iş-zaman çizelgesinde belirtilebilir)?			
Sürüm denetim sistemi (Version Control System; Git, Subversion v.s.) kullanılmış mı?			
Sistemin genel testi için uygulanan metotlar ve iyileştirme süreçlerinin dökümü verilmiş mi?			
Yazılımın sızma testi yapılmış mı?			
Performans testi yapılmış mı?			
Tasarımın uygulamasında ortaya çıkan uyumsuzluklar ve aksaklıklar belirtilerek çözüm yöntemleri tartışılmış mı?			
Yapılan işlerin zorluk derecesi?	x	0-25	
Sözlü Sınav			
Yapılan sunum başarılı mı?	x	0-5	
Soruları yanıtlama yetkinliği?	x	0-20	
Devam Durumu			
Öğrenci dönem içerisindeki raporlarını düzenli olarak hazırladı mı?	x	0-5	
Diğer Maddeler			
Toplam			

DANIŞMAN (JÜRİ ADINA):

DANIŞMAN İMZASI:

