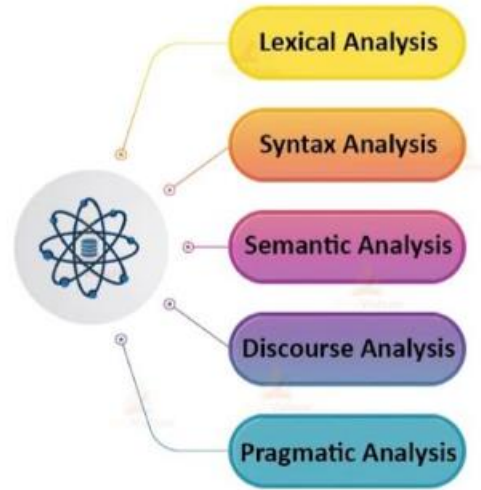


Dođal Dil İşleme Bilgi Düzeyleri-1



Prof.Dr. Banu Diri

Dođal Dil İşleme Bilgi Düzeyleri

6 ana başlık altında toplanabilir:

➤ **Sesbilimi (Phonology)**

Dilin ses özelliklerini inceleyen

➤ **Biçimbilim Bilgisi (Morphology)**

Sözcüklerin yapısını inceleyen

➤ **Sözdizim Bilgisi (Sentaks)**

Sözcüklerin yapısal ilişkilerini inceleyen

➤ **Anlamsal Bilgi (Semantic)**

Sözcük, söz öbekleri ve tümcelerin anlamlarını inceleyen

➤ **Söylem Bilgisi (Discourse)**

Farklı türdeki metinlerin yapısını inceler (Mayo ve bikini ile denize girmek yasaktır)

➤ **Kullanım Bilgisi (Pragmatics)**

Dış dünyadan (metin içeriğinin dışından) gelen bilgileri inceler (Bu yakınlarda metro istasyonu var mı?)

Bu bilgi düzeylerinin herbiri veriyi işlerken, kendisinden önce gelen bilgi düzeylerinin sonuçlarından yararlanır

Dođal Dil İşlemede Kullanılan Modeller

- Sonlu Durum Makineleri (Finite State Automata)
- Kural Tabanlı Sistemler (Rule Based Model)
- Mantık
- Olasılık Teorisi
- Makine Öğrenmesi
- Derin Öğrenme

Sesbilgisi(Fonetik/Phonetics)

Giriş: Akustik işaret

Çıkış: Sesbilim alfabesi

- Fonetik, konuşulurken dil, gırtlak, ses telleri, damak, dişler ve dudaklar ile çıkarılan sesleri ve bu seslerin dil ile olan ilişkilerini tanımlamak için kullanılan bir terimdir.
- Doğal dillerde anlam ayırıcı olarak kullanılan en küçük ses **fondur** (*phon*) dur. Fonetik terimi bu kökten gelmektedir.

- **Fon** kavramı evrensel değildir ve her dilde farklı seslere karşılık gelir.
- Farklı dillerdeki **fonların** tek ortak özelliği, ayırıcı temel sesler olmalarıdır.
- Sesle ifade edilen dili, yani konuşmayı kaydetmek için yazı icad edilmiştir.
- Konuşmayı yazı ile ifade etmek için ses birim veya **fonların** harflerle eşleştirilmesi gerekmektedir.
 - *Bazı dillerde, örneğin Türkçe, Fince ve Japonca'da, sesbirimler doğrudan harflere karşılık gelir ve bu tip dillere fonetik diller denir.*
 - *İngilizce, Almanca, Fransızca gibi dillerde ise Fonlar harflere karşılık gelmezler. Bu yaklaşımın yerine uluslararası geçerliliği olan fonetik bir alfabe, ses birimlerini ifade etmek için kullanılır.*

Fonemlerin seslendirilmesiyle ses birimleri (phon) oluşur.

Fonetik Alfabe

Harf	NATO	Türkiye	Uluslararası
A	Alpha	Ankara	Amsterdam
B	Bravo	Bursa	Baltimore
C	Charlie	Ceyhan	Casablanca
Ç	-	Çankırı	-
D	Delta	Denizli	Danemark
E	Echo	Edirne	Edison
F	Foxtrot	Fatsa	Florida
G	Golf	Giresun	Golf
H	Hotel	Hopa	Havana
I	India	Isparta	Jerusalem
İ	-	İskenderun	-
J	Juliet	Jale	Jerusalem
K	Kilo	Kayseri	Madagaskar
L	Lima	Lüleburgaz	Liverpool
M	Mike	Manisa	Madagaskar
N	November	Nazilli	New York
O	Oscar	Ordu	Oslo

Harf	NATO	Türkiye	Uluslararası
Ö	-	Ödemiş	-
P	Papa	Pazar	Paris
Q	Quebec	-	Quebec
R	Romeo	Rize	Roma
S	Sierra	Samsun	Santiago
Ş	-	Şarköy	-
T	Tango	Trabzon	Tripoli
U	Uniform	Urfa	Upsala
Ü	-	Ünye	-
V	Victor	Van	Valencia
W	Whiskey	-	Washington
X	X-ray	-	Xanthippe
Y	Yankee	Yalova	Yokohama
Z	Zulu	Zonguldak	Zurih

Ses Bilim (Fonoloji/Phonology)

Giriş: Fonetik alfabe içerisindeki sesler

Çıkış: Sıralı fonemler

Ses bilimi, (fonoloji) belirli dil seslerini (fonemler) inceleyen bir bilim dalıdır. Ses Bilimi dil içindeki seslerin işlevlerini inceler.

Biçimbilim (Morfoloji/Morphology)

- Dil bilime terim olarak 1859 yılında August Schleicher tarafından kazandırılmıştır.
- Morfoloji, dilde biçimi oluşturan öğelerin türlerini tanımlamak ve özetle dil bilgisi kuralları denen biçimsel öğelerin sınıflandırmasını yapmaktır.

Sözcük Nedir ? İki ucuna boşluk verilerek yazılan öge

Yazımsal sözcük *orthographic word* tanımına göre

Çocuklar sabahtan beri evin önünde top oynuyor.

Ayşe gelmişti Ayşe gelmiş idi **7 sözcük**
arabayla araba ile

Yazımsal Sözcüğün **karşıtı** Sesbilimsel Sözcüktür (*phonological word*)

Sözcüğün ünlü uyumları, vurgu gibi sesbilimsel ölçütlere dayandırılması

Çocuklar sabahtan beri evin önünde top oynuyor.

Chatgpt cevabı : Çocuklar sabahtan beri evin önünde top oynuyor.

Sözdizimsel (Sentaks/ Syntax) Analiz

Sözdizimsel analiz, sözdizimini (syntax) veya cümleyi oluşturan morfolojik öğelerin hiyerarşik kurallara uyumunu karşılaştırarak ölçümleme yapmaktadır.

Böylece, söz dizimin anlamlı olup olmadığının ölçülebilmesi için düzenleyici bir süreç gerçekleşmiş olur.

- ❖ Türkçe Ural-Altay dil grubuna girer.
- ❖ Türkçe **bitişken** (*agglutinative*) bir dildir. (Fince ve Macarca)
- ❖ Kelimeler, bir kök sözcüğe eklenen **biçimbirim**lerden (*morpheme*) oluşur. Ancak, bu ekleme işlemi sırasında ünlü uyumu, ünsüz değişmesi, ünlü ve ünsüz düşmesi gibi değişiklikler gerekebilir.
- ❖ Biçimbirimler eklendikleri kök ve gövdenin anlamını, sözcük türünü veya söz dizimsel işleyişini değiştirebilirler.
- ❖ Türkçe'de bu şekilde oluşturulan bir kelime bazen bir başka dilde bir cümle ile ifade edilebilir.
- ❖ Kök : yer (yerleştirilemeyeceklerindedir)

Yer+leş+ti+ri+le+me+ye+cek+le+rin+den+dir

❖ Türkçe'de cümleler en genel şekliyle özne, nesne ve yüklem'den oluşur.

❖ Cümleye eklenmek istenen anlamlar arttıkça cümleler, özne, yer tamlayıcısı, zarf tamlayıcısı, nesne ve yüklem gibi bileşenleri içerir.

❖ Cümlenin anlamını kuvvetlendiren cümle dışı bileşenler de (bağlaç, edat, vb.) cümlede bulunabilir (ile, için, ama, çünkü, vs.).

❖ Türkçe'de özne ile yüklem cümlenin temel bileşenleridir ve genelde tüm cümlelerde yer alırlar.

❖ Yer tamlayıcısı, zarf tamlayıcısı, nesne gibi bileşenler bazı cümlelerde yer almayabilirler veya bazı cümlelerde sadece biri, bazılarında sadece ikisi bulunabilir. Bu bileşenlerin cümle içindeki sıralanışları da değişebilir.

Bilgisayar ile dođal dilin modellenmesinde anlamsal analizden önce, kelimelerden oluşturulan yapının cümle olup olmadığının test edilmesi gerekir.

Bu işlem sentaktik eşleştirme işleminde anlamsız eşleşmelerin önlenmesine yardımcı olur.

Anlambilimsel (Semantic) Analiz

- ❖ **Anlambilimsel analiz**, sözdizimini oluşturan [morfolojik](#) öğelerin ayrılması yani, sözdizimsel analiz ile anlam taşıyan kelimelerin sınıflandırılması işleminden sonra gelen, anlamlandırma veya anlama sürecidir.
- ❖ Bu süreçte anlam taşıyan kelimelerin, ekler ve cümle hiyerarşisi içindeki konumlarınının saptanması sayesinde birbirleri ile ilişkileri kurulabilir.
- ❖ Bu ilişkiler **anlam çıkarma**, **fikir yürütme** gibi ileri seviye bilişsel fonksiyonların oluşturulmasında ham bilgi olarak kullanılmaktadır.

Söylem (Discourse) Analiz

«**söylem**», tek bir cümlenin ötesinde insanlar arasındaki sözlü veya yazılı iletişimi ifade eder.

- Söylem sadece bir dil değildir.

"Dil", tüm dilbilimsel ve sembolik birimleri içerebilir ve dil çalışmaları kelimelerin bireysel anlamlarına odaklanabilir.

- Söylem, bunun ötesine geçer ve dilin bağlam içinde ilettiği genel anlamlara bakar. Buradaki "bağlam", söylemin sosyal, kültürel, siyasi ve tarihsel arka planını ifade eder ve dil aracılığıyla ifade edilen altta yatan anlamları anlamak için bunun dikkate alınması önemlidir.
- Söylem analizi, bir veri kümesinde sunulan dil kullanılarak içerikten anlam çıkarmak için kullanılır.

- Bir tür güç veya eşitsizlik

(örneğin, varlıklı bireylerin daha az varlıklı olanlarla etkileşim şekli)

- İnsanların belirli bir bağlamda nasıl iletişim kurduğu

(örneğin, meslektaşlarla bir sosyal ortamda nasıl iletişim kurduğu ile bir kurul toplantısı arasındaki fark)

- İdeoloji ve fikirlerin (değerler ve inançlar gibi) dil kullanılarak nasıl paylaşıldığı

(örneğin, siyasi konuşmalarda olduğu gibi)

- İletişimin nasıl sosyal hedeflere ulaşmak için kullanıldığı

(örneğin, bir dostluğu sürdürmek veya çatışmayı yönetmek)

Söylem analizi, toplumsal sorunları değerlendirmek için güçlü bir araç olabilir, ayrıca güç ve güç dengeleri hakkında da bize bilgi verir.

Kullanım (Pragmatic) Analiz

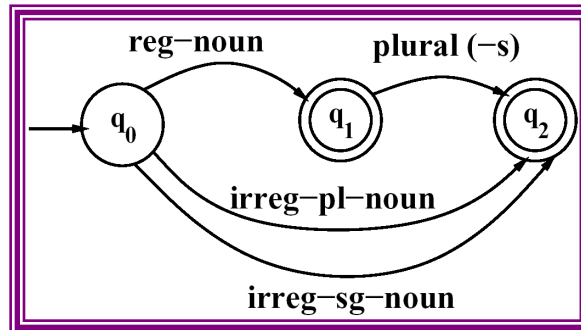
- Cümlenin gerçek anlamını çıkarmak için gerçek dünya bilgisini veya bağlamı kullanır.
- İletişimdeki söylemlerin yorumlanmasıyla ilgilenen dilbiliminin bir alt dalıdır.
- Pragmatik Analiz, konuşmacının ve yazarın niyetleri ile birlikte, bağlamı oluşturmak için de ek bilgiler kullanır.

**"Bana bir bardak su ver" bir emir mi ? yoksa bir istek mi ?
Bağlama sahip olmadan söyleyemeyiz.**



" Öğretmenin saati soorduğu" anlamına gelmez. Cümlenin semantik anlamı bu olsa bile, öğretmen, "Neden bu kadar geç kaldın?" anlamında sormaktadır. Bu da cümlenin pragmatik anlamıdır.

Biçimbirimsel Analiz



Biçimbirimsel Çözümleme (Morphological Analysis)

- ❖ Biçimbirimsel (Morphological) Analiz, bir metin içerisinde yer alan her kelimenin bilgisayar tarafından otomatik olarak çözümlenmesi işlemidir.
- ❖ Kelimelerin, kök ve eklerine ayrıştırılıp, görevlerinin belirlenmesi ile ilgilenir.

Kök ve ekler morphem ismini de verebiliriz.

Kelimenin en küçük anlamlı birimi **morpheme**

fox (tilki) → 1 morpheme

cats (kediler) → 2 morpheme (*cat* ve *-s* çoğul eki)

Örnek: çiçekler → *çiçek* (kök) + *ler* (çoğul eki) [*'çiçek'* ve *'ler'* birer morphemdir]

Morphem 'ler

- gövdeler (stem) (kelimenin ana morpheme dir ve kelimeye anlamını verir)
- ekler (affixes) (kelimenin anlamı şekil değiştirir)

Affix'ler

- prefix (gövdenin önünde yer alır) (impossible, natamam)
- suffixes (gövdeyi takip eder) (eats, yemekler)
- infixes (gövdenin içerisine eklenir-Filipince-Tagalog)
- circumfixes (hem gövdenin önünde hem de arkasında yer alır)
(almancada → sagen (verb- to say) gesagst (past participle- said))

Morphem ler birleştirilerek yeni kelimeler yaratılır.
Konuşma ve dil işlemede önemlidir.

- ✓ Inflection (Çekim eki) : -s, -ed (ing), -lar, -ler (tr)
- ✓ Derivation (Yapım eki) : computerize (verb)
computerization (isim)
göz, gözlük
- ✓ Compounding (Birleştirme) : Birden fazla gövde halindeki kelimeyi birleştirir.
dog +house → doghouse
pazar + yer → pazaryer+i
- ✓ Cliticization : I have → I've
an apple
*l'*opera
-de, -ki, -mi

Bir kelime bir veya birden fazla affix (ek) alabilir.

rewrites (prefix *-re*, stem *-write*, suffix *-s*)

unbelievably (stem *-believe*, affix *-un*, *able*, *ly*)

- İngilizce 4 veya 5 ten fazla ek almaz.
- Türkçe gibi *agglutinative –sondan ekli* diller 10 dan fazla da ek alabilirler.

uygarlaştıramadıklarımızdanmışsınızcasına

uygar +*laş* +*tır* +*ama* +*dık* +*lar* +*ımız* +*dan* +*muş* +*sınız* +*casına*

civilized +BEC +CAUS +NABL +PART +PL +P1PL +ABL +PAST +2PL +AsIf

“(behaving) as if you are among those whom we could not civilize”

The various pieces of this word (the **morphemes**) have these meanings:

+BEC “become”

+CAUS the causative verb marker (‘cause to X’)

+NAB “not able”

+PART past participle form

+P1PL 1st person pl possessive agreement

+2PL 2nd person pl

+ABL ablative (from/among) case marker

+AsIf derivationally forms an adverb from a finite verb

Örnek : İngilizce için : “araba” → tekil (*car*) ve çoğul (*cars*)

kelimenin bütün halleri bir listede tutularak kelimenin çözümlenmesi yapılabilir

“*You had been in the car.*” (arabadaymışsınız) cümlesindeki “car” kelimesi kolayca çözümlenir.

Türkçe’de durum nedir?

araba	arabalar	arabam	araban
arabası	arabamız	arabanız	arabaları
arabalarım	arabaların	arabaları	arabalarım
arabalarınız	arabayı	arabada	arabalarda
arabadayım	arabadası	arabasında	arabasınd
arabadaymı	arabasınd	arabalardak	arabandan

- Türkçe karmaşık bir biçimbirimsel yapıya sahiptir
- Kelimenin olası tüm biçimlerini listelemek mümkün değildir
- Türkçe’de tek bir kelime, aldığı eklerle bir çok anlamı üzerinde taşıyabilir

Arabadaymışsınız → *you had been in the car*

araba (isim) → *the car*

+da (İsmin –de hali) → *in*

+ymış (-miş ekfiil eki ile eyleme dönüşme) → *had been*

+sınız (2. çoğul kişi eki) → *you*

Biçimbirimsel ayrıştırma sonucunda:

- Her kelime için birden fazla olası çözümleme olabilir
- Hangisinin doğru olduğuna karar vermek biçimbirimsel çözümleyicinin görevi değildir

“kalem”

1. Yalın haldeki “kalem”

2. ‘-m’ 1.tekil kişi iyelik eki almış “kale”

Hangisinin doğru olduğuna bağlama bakarak karar verebiliriz.

“Bu kalem yazmıyor.” (1)

“Benim kalem yıkıldı.” (2)

Örnek: çocukları

(Sevgi'nin çocukları Ayşe ve Mehmet geldiler.)

çocuk+İsim+ Çoğul+ 3.tekil kişi iyelik

(Yeni gelen çocukları gördünüz mü?)

çocuk+İsim+ Çoğul+ -i hali

(Ayşe ile Mehmet'in çocukları Gökhan ile Sevgi'dir.)

çocuk+İsim+ Çoğul+ 3. çoğul kişi iyelik

(Ayşe'nin çocukları Gökhan ile Sevgi'dir.)

çocuk+İsim+ Tekil+ 3. çoğul kişi iyelik

Nerelerde ihtiya duyulur

- Internette arama yaparken
- Yazım denetiminde
- Bul deęiřtir (“oęul” → “kız” ya oęlum yazıldıysa, bina→ev binadan→evden)
- Makine evirisinde (“arabaymıřsınız”→ “You had been in the car.”)

Çekim eki Çözümleme

İngilizce çekim ekleri

noun, verb, adjective uygulanır.

Affix → plural, possessive (sahiplik) (books, boxes, children's)

	Regular nouns		Irregular nouns	
Singular	Cat	Thrush	Mouse	Ox
Plural	Cats	Thrushes	Mice	Oxen

Morphological class	Regularly Inflected Verbs			
stem	walk	merge	try	map
-s form	walks	merges	tries	maps
-ing participle	walking	merging	trying	mapping
past form or <i>-ed</i> participle	walked	merged	tried	mapped

Morphological class	Irregularly Inflected Verbs	
stem	eat	catch
-s form	eats	catches
<i>-ing</i> participle	eating	catching
preterite (geçmiş z.)	ate	caught
past participle	eaten	caught

Türetme eki çözümleme

- ✓ İngilizcede kelime türetme oldukça karışık bir yapıya sahiptir.
- ✓ İngilizcede en çok kullanılan türetme şekli fiil ve sıfatlardan isimlerin türetilmesidir. Bu işleme *nominalization* denir.

Suffix	Base Verb/Adjective	Derived Noun
<i>-ation</i>	computerize (V)	computerization
<i>-ee</i>	appoint (V)	appointee
<i>-er</i>	kill (V)	killer
<i>-ness</i>	fuzzy (A)	fuzziness

Suffix	Base Verb/Noun	Derived Adjective
<i>-al</i>	computation (N)	computational
<i>-able</i>	embrace(V) <small>kucaklamak</small>	embraceable
<i>-less</i>	clue (N)	clueless

Cliticization

- proclitics (vurgu bakımından kendisinden sonra gelen kelimeye bağlıdır.
an apple)
- enclitics (hem kitap hem *de* kalem aldı)

Full Form	Clitic	Full Form	Clitic
am	`m	have	`ve
are	`re	has	`s
is	`s	had	`d
will	`ll	would	`d

➤ Biçimbirimsel analizde kelimelerin sadece kök ve eklerinin çözümü yapılmaz. Aynı zamanda kelimelerin isim, fiil, sıfat, zarf, edat gibi tipleri de belirlenir.

➤ Biçimbirimsel Analiz'in çıktısı, Sözdizimsel (Syntax) Analizde kullanılır.

➤ Biçimbirimsel Analiz için teoride tüm kelimeler ve onların olası formları bir sözlükte tutularak gerektiğinde erişilebilir. Ancak, Türkçe ve Fince gibi sondan eklemeli dillerde bir kelimenin kökünden çok sayıda kelime türetildiğinden biçimbirimsel çözümlemenin yapılması şarttır.

Biçimbirimsel Analizde 3 temel elemana ihtiyaç vardır.

- Sözlük
- İmla Kuralları
- Biçimbirimsel Kurallar

❖ **Lexicon (Sözlük)** : Dildeki tüm gövde ve eklerin tutulduğu bir yapıdır. Ancak Türkçe gibi sondan eklemeli bir dilde tüm kelimelerin sözlükte yer alması imkansızdır. Bu yüzden kelimedeki değişimler ayırıcı işaretler ile sözlükte tutulmalıdır.

❖ **Morphotactics Rules (Biçimbirimsel Kurallar)** : Morphemlerin eklenme sırasının kuralları

❖ **Orthographic Rules (İmla Kuralları)** : Kelimeler içerisinde meydana gelen değişimleri tanımlayan kurallardır. Bazen morfemlerin arka arkaya bağlanma durumlarında morfemler değişime uğrayabilir.

- ünlü uyumu,
- ünlü düşmesi (oğul→oğlu),
- ünlü daralması (anla+yor→anlıyor),
- ünsüz sertleşmesi (meslek +daş→meslektaş),
- ünsüz yumuşaması (ağaç+a→ağaca),
- ünsüz türemesi (his+i→hissi)

gibi.

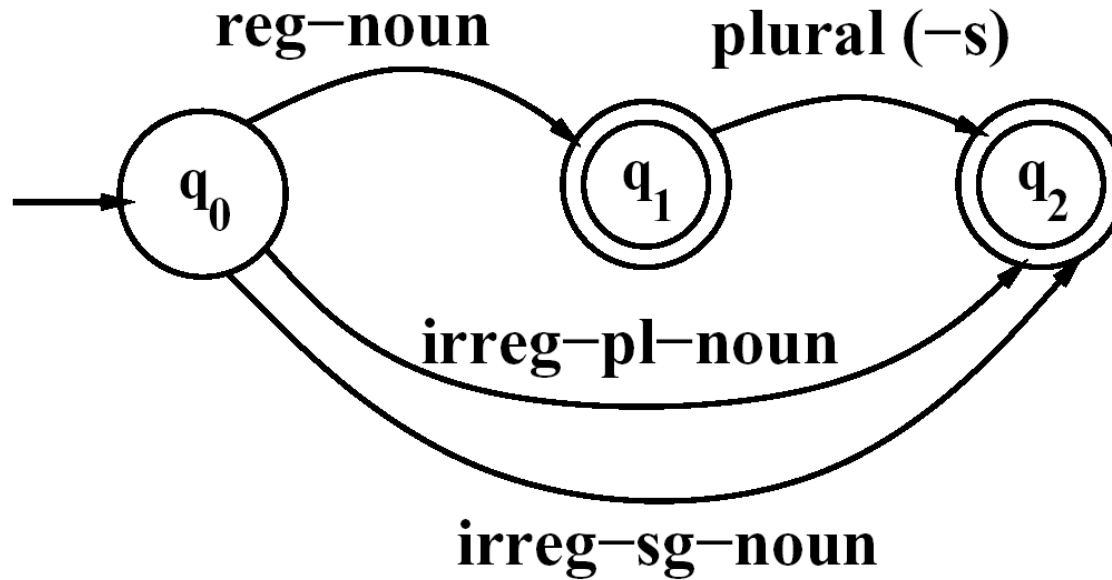
Örnek: İngilizcede bir kelimenin nasıl çoğul yazılacağını söyler.

(kelime sonundaki *-y* düşer, yerine *-i* gelir ve arkasına *-es* eklenir)

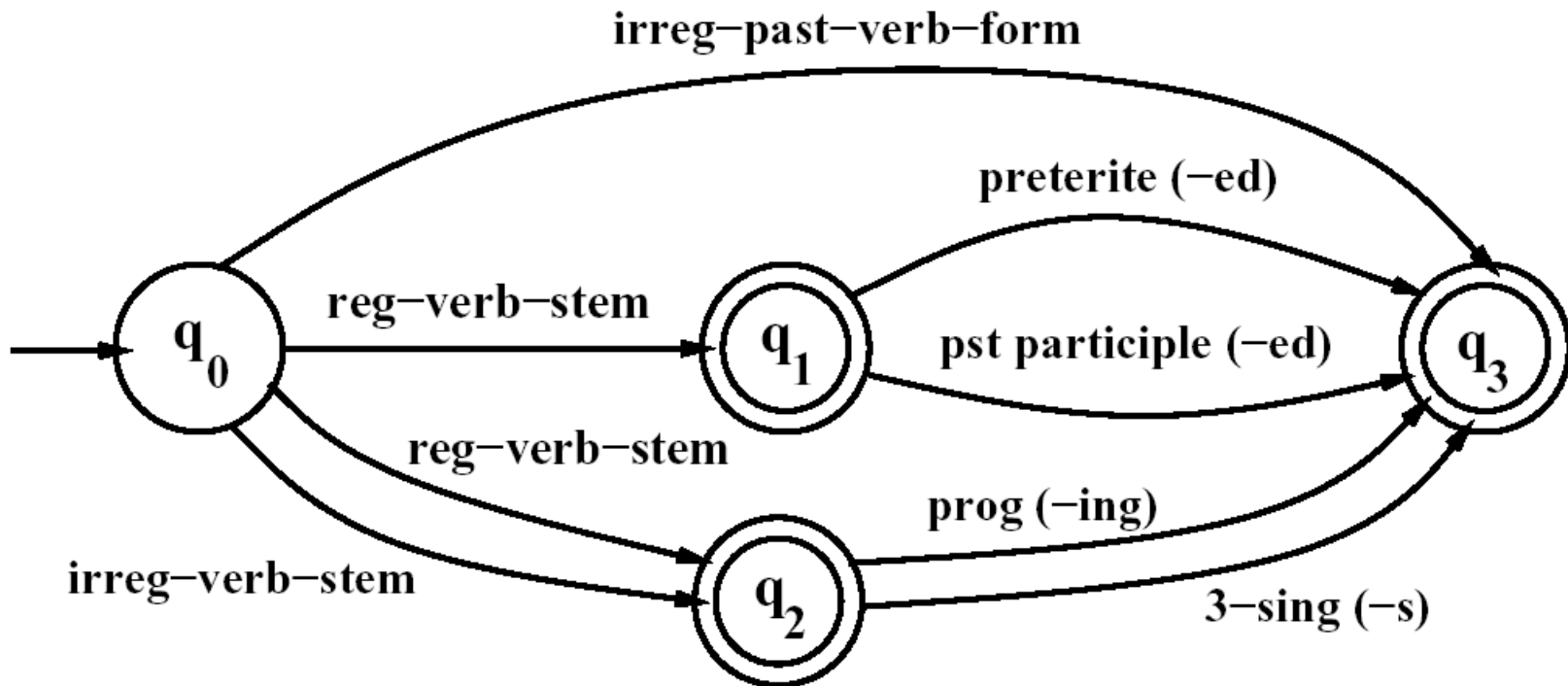
Parsing English morphology

Input	Morphological parsed output
cats	cat +N +PL
cat	cat +N +SG
cities	city +N +PL
geese	goose +N +PL
goose	(goose +N +SG) or (goose +V)
gooses	goose +V +3SG
merging	merge +V +PRES-PART
caught	(caught +V +PAST-PART) or (catch +V +PAST)

Construction of a Finite State Lexicon



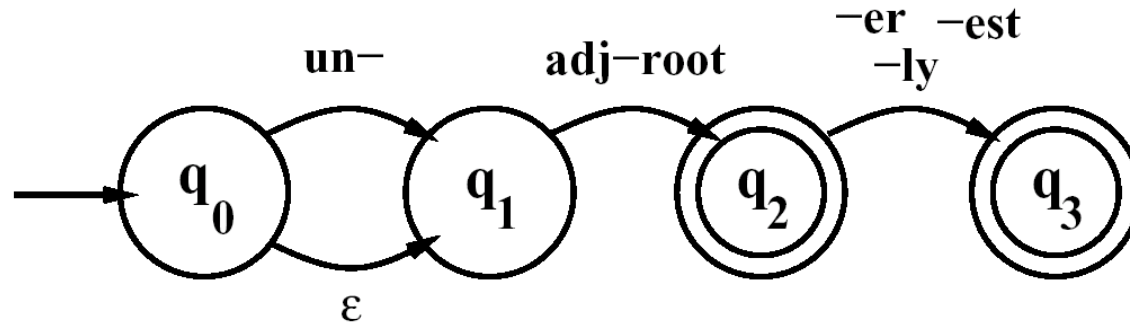
reg-noun	irreg-pl-noun	irreg-sg-noun	plural
fox cat armadillo	geese sheep mice	goose sheep mouse	-s



Reg-verb-stem	Irreg-verb-stem	Irreg-past-verb	past	Past-part	Pres-part	3s g
walk fry talk impeach	cut speak sing sang spoken	caught ate eaten	-ed	-ed	-ing	-s

İngilizcede kelimeye yapılan türetme işleminin çekim işlemine göre kontrolü daha zordur.

İngilizcede sıfatlar için oluşturulmuş basit bir yapı.



Morphology #1

big, bigger, biggest

cool, cooler, coolest, coolly

red, redder, reddest

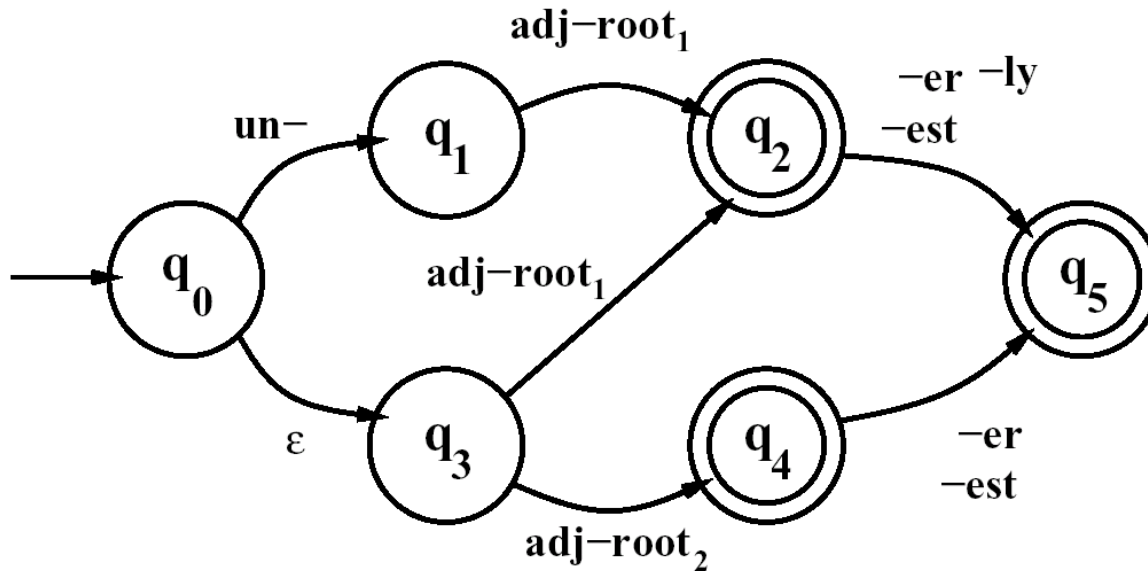
clear, clearer, clearest, clearly, unclear, unclearly

happy, happier, happiest, happily

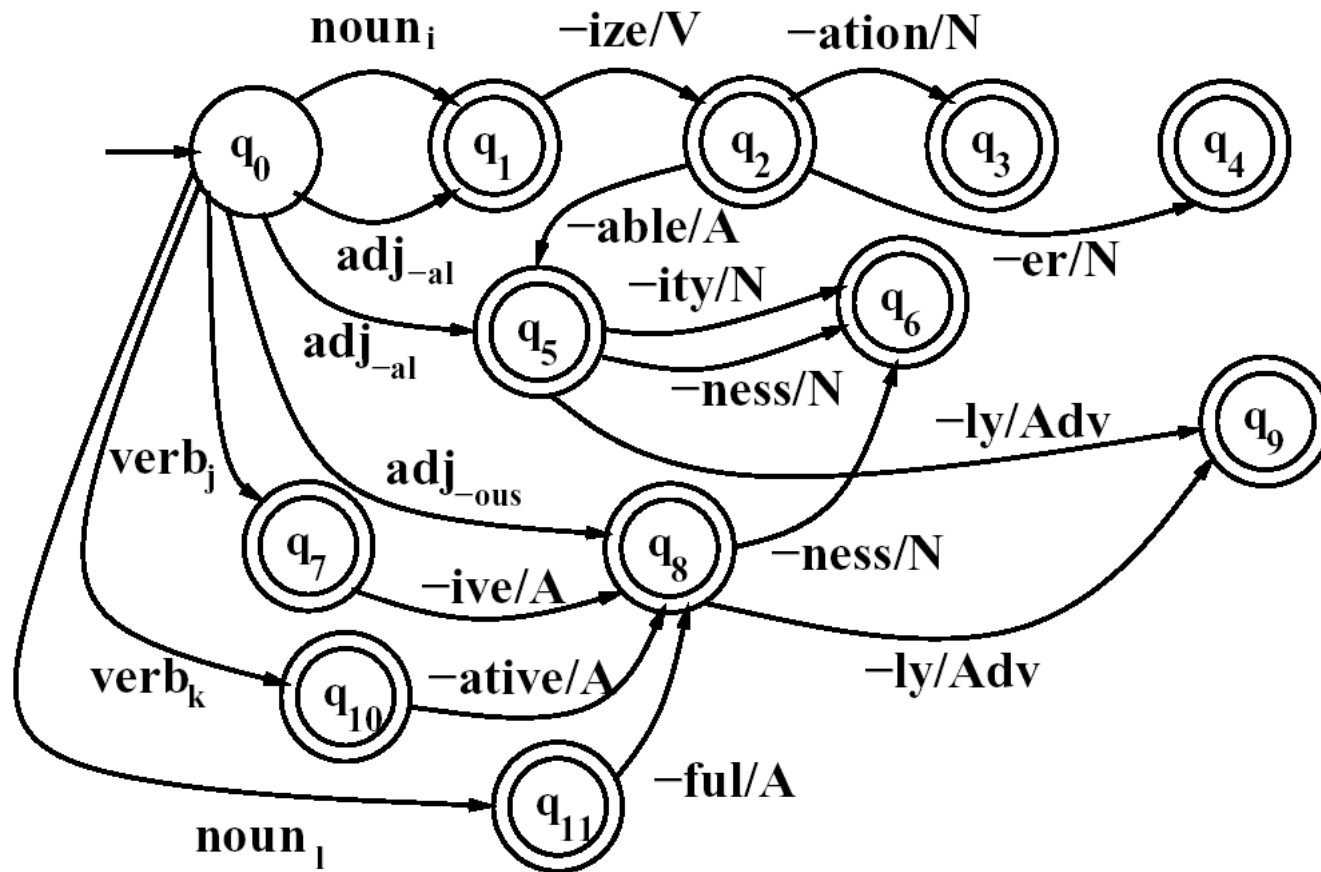
unhappy, unhappier, unhappiest, unhappily

real, unreal, really

- FSA#1 listelenen tüm sıfatları tanır ancak hatalı yazılmış *unbig*, *redly*, ve *realest* kelimelerini de tanır.
- #1 üzerinde değişiklik yaparak #2 oluşturalım.
- root1 \rightarrow *-un*, *-ly* ile kullanılan sıfatlardan
- root2 \rightarrow diğerlerinden oluşan bir sözlük olsun.



An FSA for a fragment of English adjective Morphology #2



An FSA for another fragment of English **derivational morphology**

Kelimenin Kökünün Bulunması. Kelimenin kökünü bulmak için bir sözlüğe ihtiyaç vardır. Bu sözlük, içerisinde ilgili dilde yer alan :

- tüm kök ve gövdeleri,
- bunların tür bilgileri,
- kelimedede oluşabilecek ses düşmeleri gibi değişimler hakkında bilgiler de tutulmalıdır.

Kök Belirleme

Türkçe bir kelimenin biçimbirimsel ayrıştırması yapılmadan önce kelimenin kökünün doğru olarak belirlenmesi gerekir. Öncelikle Türkçe'deki kelimeleri ve dildeki düzensiz yapıları içerisinde barındıran bir sözlük oluşturulur. Bu sözlükte, her kelimenin yapısal özelliklerini gösteren bayrak alanları bulunmalıdır.

1. Maksimum eşleme (maximal match) algoritması ile kelimenin kökü aranır.
2. Önce kelimenin tamamı sözlükte aranır. Kelime sözlükte bulunuyorsa ek almamış olarak belirlenir ve herhangi bir ayrıştırma yapılmaz.
3. Kelime bulunmaz ise her seferinde kelimenin sağından bir harf atılarak, kalan kısım sözlükte tekrar aranır.

(Bazı çalışmalarda bu işleme soldan başlanır)

4. Bu işlem herhangi bir kök bulununcaya kadar tekrar edilir. Kelime tek harf olarak kaldığında, kelimenin hatalı yazılmış olduğuna karar verilir.

5. Bu algoritma ile her zaman doğru kök bulunmayabilir. Kök bulunduktan sonra kelimenin geri kalan kısmı kurallara uygun bir şekilde ayrıştırılamıyorsa kelimenin yanlış bulunduğuna karar verilir ve yeni bir kök aranır.

Örnek:

Yazıldın kelimesine bakalım. Sağdan harf atarak gittiğimizde *yazı* kelimesi kök olarak bulunur. Fakat kelimenin geri kalanı doğru olarak ayrıştırılamaz. *Yazı* kökü geçersiz sayılır ve yeni kök aranır. Sonra *yaz* kelimesi kök olarak bulunur.

6. Türkçe’de bazen kelime kökleri aldıkları eklere göre değişebilir. Bu durum kelimenin kökünün bulunmasını zorlaştırır. Değişen bu köklerin bulunabilmesi için oluşturulmuş olan sözlükte ilgili kelimenin yanına bir bayrak yardımıyla işaret konulur.

Örnek: *oğlumuz* kelimenin kökü *oğul* dur.

Kök hece düşmesine uğramış *oğl* olmuştur.

Oğl kelimesi sözlükte bulunamaz. Kullanılan algoritmada değişiklikler yapılarak hece düşmesine uğrayan kelimelere bir işaret konulur. Son iki harfi sessiz olan ve bir sonraki harfi sesli olan kelimelerde, son iki sessiz harfin arasına bir sonraki sesli harf yazılarak kelime sözlükte aranır. Böyle bir kelime sözlükte bulunur ve hece düşmesine uğrayacağı işaretlenmiş ise bu kelime kök olarak bulunur.

Oğulumuz yazılan kelime de oğul kök olarak bulunur ancak hece düşmesine uğramadığı için yanlış yazılmış kabul edilir.

7. Türkçe’de kökte oluşan bir başka değişimde hece yumuşamasıdır.

b,c,d, g/ğ → p, ç, t, k

Örnek: *bacağım* kelimesinin asıl kökü *bacak* ‘tır. *Bacağ* kelimesi sözlükte bulunamayacaktır. Hece yumuşaması kuralını uyguladığımızda *bacağ*, *bacak* olarak değişecek ve sözlükte hece yumuşaması bayrağı ile birlikte bulunacağından kök olarak kabul edilecektir.

8. Kelimenin kökünde oluşabilecek diğer değişiklikler de benzer çözümlerle algoritmaya eklenerek çözülebilir.

9. Kök belirleme işleminde bazen çok fazla işlem yapılabilir.

okula → okul

aldığımız → aldığımız, aldığım, aldığı, aldığı, aldık, aldı, ald, alıd, alıt, alt, al

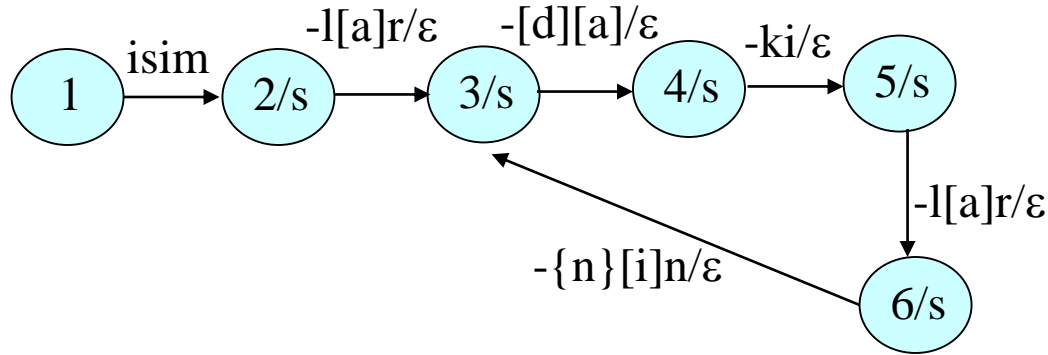
Örnek

Vasif
Nabiyey

İsim: isim kökü + çoğul eki + iyelik eki + durum eki + ki eki

Fiil: fiil kökü + çatı eki + olumsuzluk eki + bileşik fiil eki + ana zaman eki + soru eki + yardımcı zaman eki + kişi eki

Türkçe için kısıtlı sayıdaki isim biçimbirimlerini tanıyabilen bir yapı



akıllardakilerin--> isim + l[a]r + [d][a] + ki + l[a]r + {n}[1]n

akıl + lar + da + ki + ler + in

evlerdeki--> isim + l[a]r + [d][a] + ki + l[a]r + [d][a]

ev + ler + de + ki + ler + de

Biçimbirim testler

Kelimenin köküne ulaştıktan sonra kelimenin geri kalanı ek dizisi olarak düşünülür, eklerin dizilişi ve ses uyumlarının Türkçe dil yapısına uygun olup olmadığı kontrol edilir.

Kaynaştırma harfleri ve harf düşmeleri de bu kısımda incelenir.

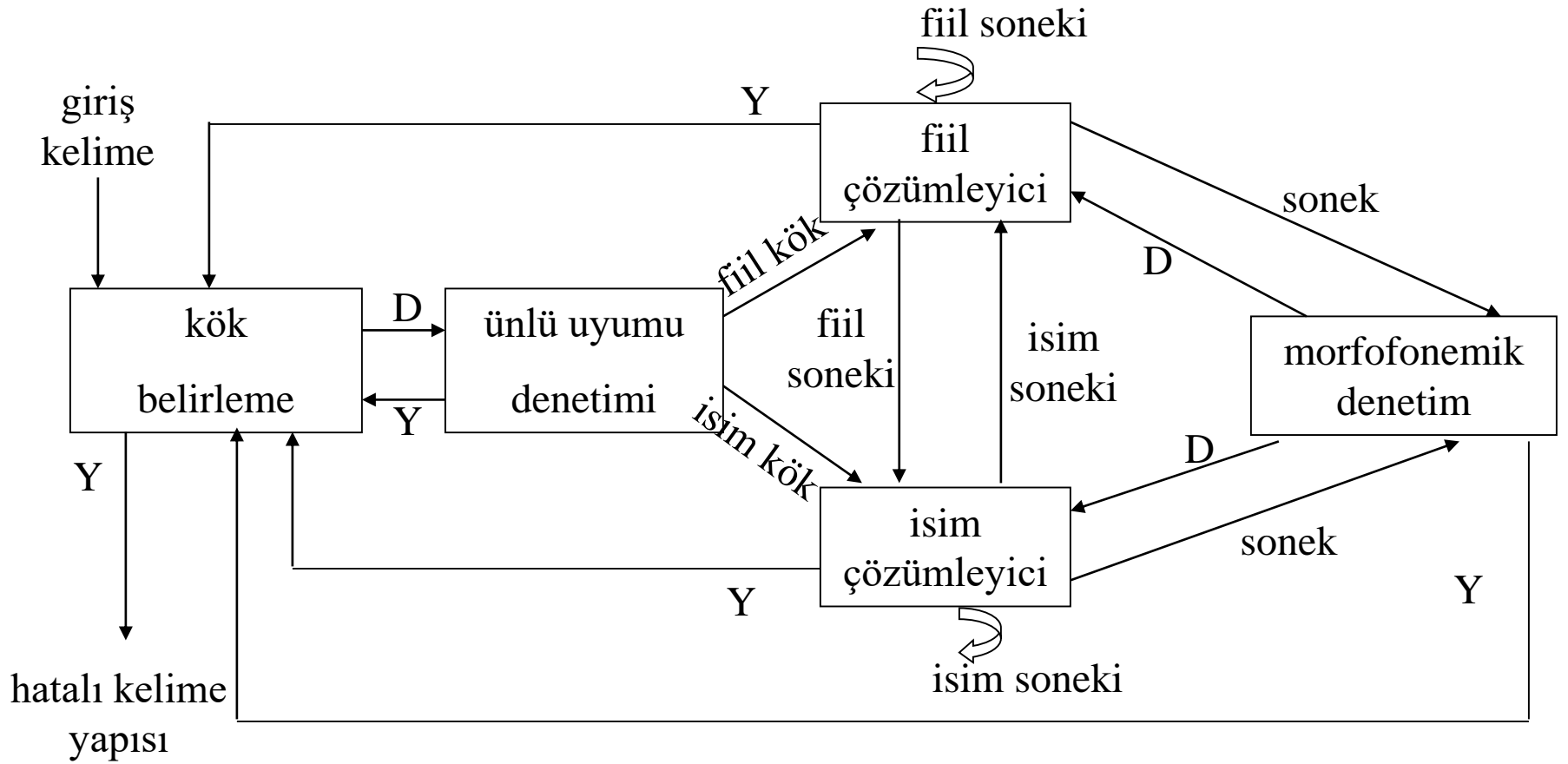
1. İlk test olarak ünlü uyumu testi yapılır.
2. Kelimenin aldığı eklerde olabilecek değişikliklerin doğruluğu kontrol edilir. Bu testler kelime eklerine ayrıştırıldıktan sonra yapılır.

Örnek: *yapdıkça, yaptığça, yaptıkca, yaptığca* ve *yaptıkça* kelimelerinin hepsi *yap* köküne sahiptir.

Hepsi ünlü uyumundan geçer.

yap+dık+ca yapısına göre ayrıştırılır. Fakat sadece *yaptıkça* doğru yazılmıştır. Yap kökünden sonra gelecek *-dık* ekinin sert sessiz uyumundan dolayı *-tik* olarak değişmesi gerekmektedir.

Aynı sert sessiz uyumdan dolayı da *-ca, -ça* olarak değişir.



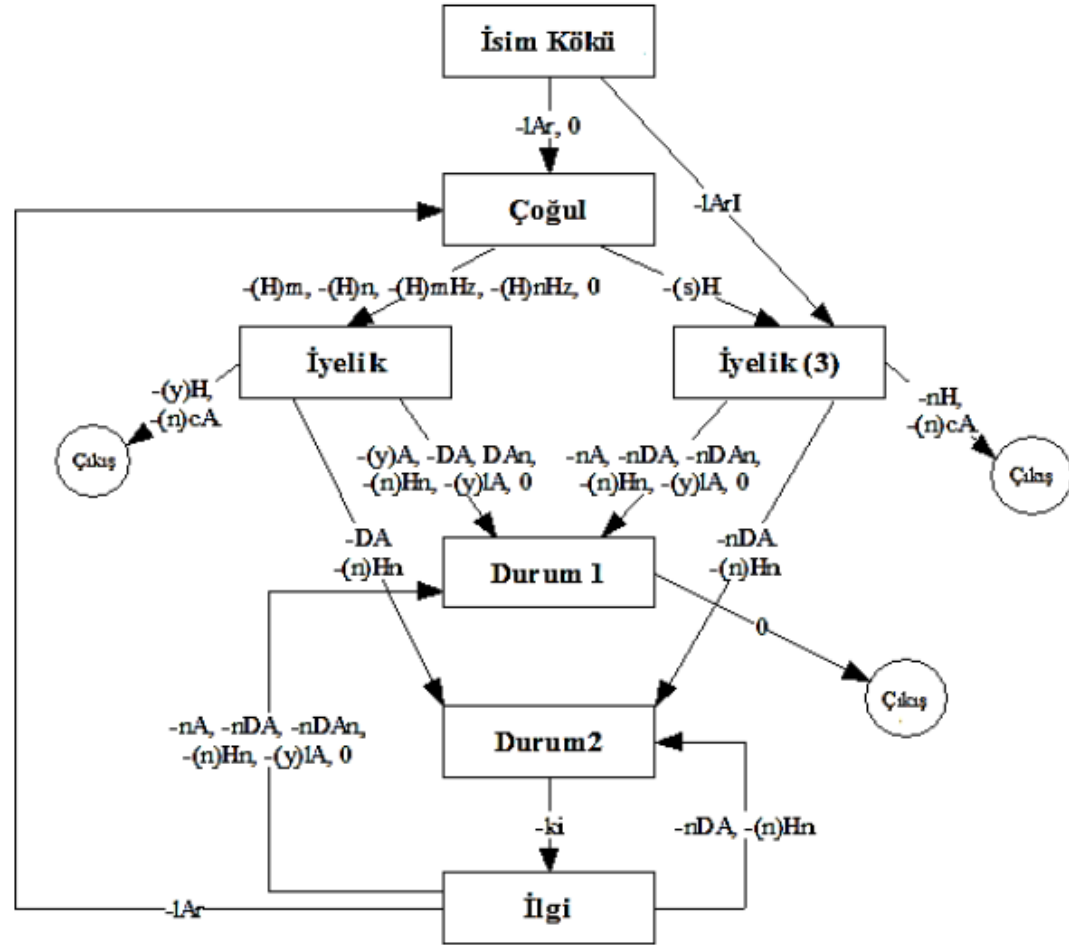
Türkçe'nin morfolojik analizi için *K.Oflazer* ve *C.Bozşahin* tarafından önerilen bir yapı bu alandaki çalışmalarda kullanılmaktadır.

- Türkçede ekler, kelimelere bazı kurallara göre eklenir
- Kelime gövdesine birden fazla ek eklemek istersek belli bir sıraya göre eklenmesi gerekir

- ✓ İsim Çekim Ekleri Kümesi
- ✓ Ek-Eylem Ekleri Kümesi
- ✓ Eylem Zaman Ekleri Kümesi
- ✓ Eylem Çekim Ekleri Kümesi

İsim Çekim Ekleri Kümesi

Ek 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
9	-nH	-i hali (3.t.k. iyelikten sonra)	anne-si-ni sonra)
10	-(n)Hn	tamlama	anne-nin
11	-(y)A	-e hali	anne-ye
12	-nA	-e hali (3.t.k. iyelikten sonra)	anne-si-ne sonra)
13	-DA	-de hali	anne-de
14	-nDA	-de hali (3.t.k. iyelikten sonra)	anne-si-nde sonra)
15	-DAn	-den hali	anne-den
16	-nDAn	-den hali (3.t.k. iyelikten sonra)	anne-sin-den sonra)
17	-(y)lA	birliktelik	anne-yle
18	-ki	ilgi	annem-de-ki
19	-(n)cA	görelilik	annem-ce



İsim Çekim Ekleri Soldan Sağa Sonlu Durum Makinesi

“sandıktakilerden”

sandık - ta - ki - ler - den

İsim Kökü⁰ → Çoğul⁰ → İyelik^{-DA} → Durum 2^{-ki} → İlgi

-lAr → Çoğul⁰ → İyelik^{-DAn} → Durum 1⁰ → Çıkış

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

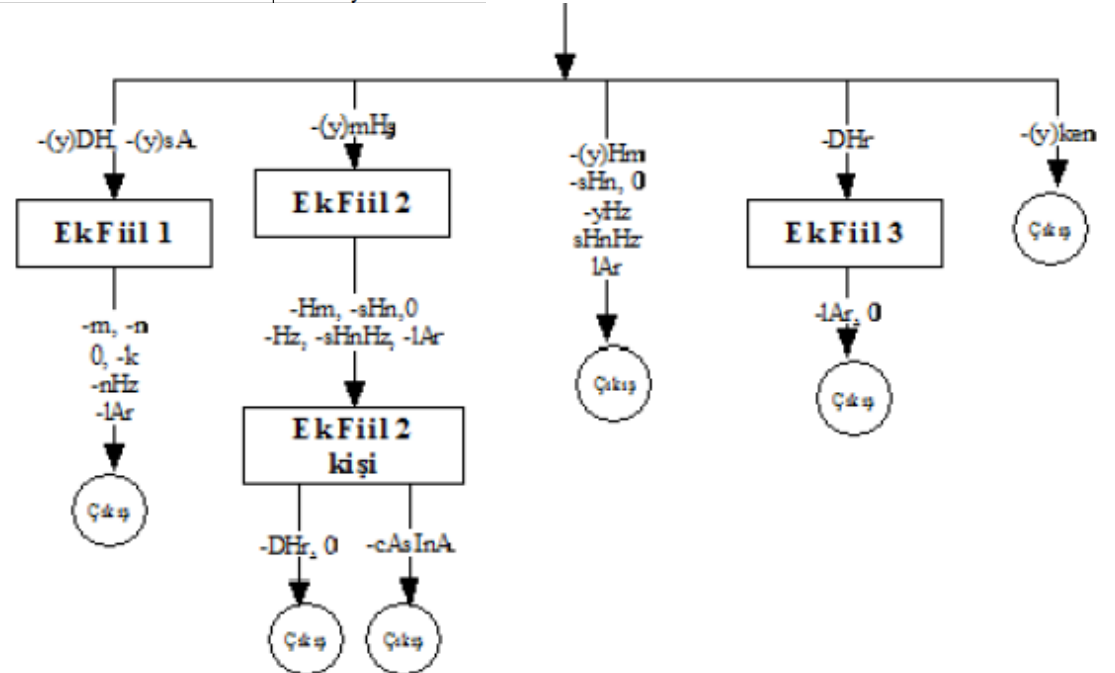
I: ı veya i harfi yerine

(): paranteziçerisindeki harf ek içinde yer almayabilir

Ek-Eylem Ekleri Kümesi isme getirilen zaman ve kişi eklerini içerir

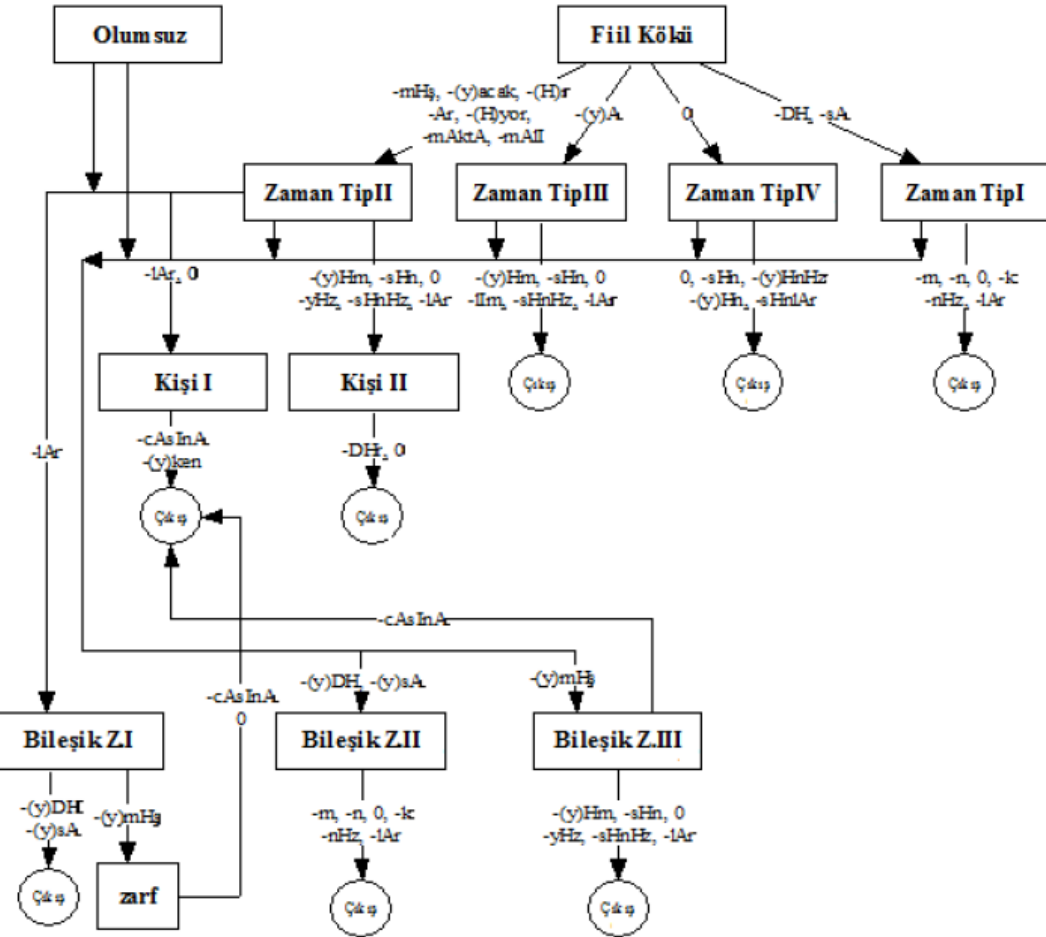
Ek No	Ek	Açıklama	Örnek
1	-(y)Hm	1. tekil kişi	akıllı-yım
2	-sHn	2. tekil kişi	akıllı-sın
3	-(y)Hz	1. çoğul kişi	akıllı-yız
4	-sHnHz	2. çoğul kişi	akıllı-sınız
5	-lAr	3. çoğul kişi	akıllı-lar
6	-m	1. tekil kişi ((y)DH ve (y)sA eklerinden sonra)	akıllı-ydı-m
7	-n	2. tekil kişi ((y)DH ve (y)sA eklerinden 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

akıllı - ymış - sın
İsim Soylu -ymış → Ekfiil2-sın → Ek fiil2 kişi → Çıkış
Sözcük



Eylem Zaman Ekleri Kümesi

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)yor	şimdiki zaman	gör-üyor
11	-mAktA	süredlilik	gör-mekte
12	-mAlI	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	-lIm	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-stin-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



gör - üyor - müş - sunuz

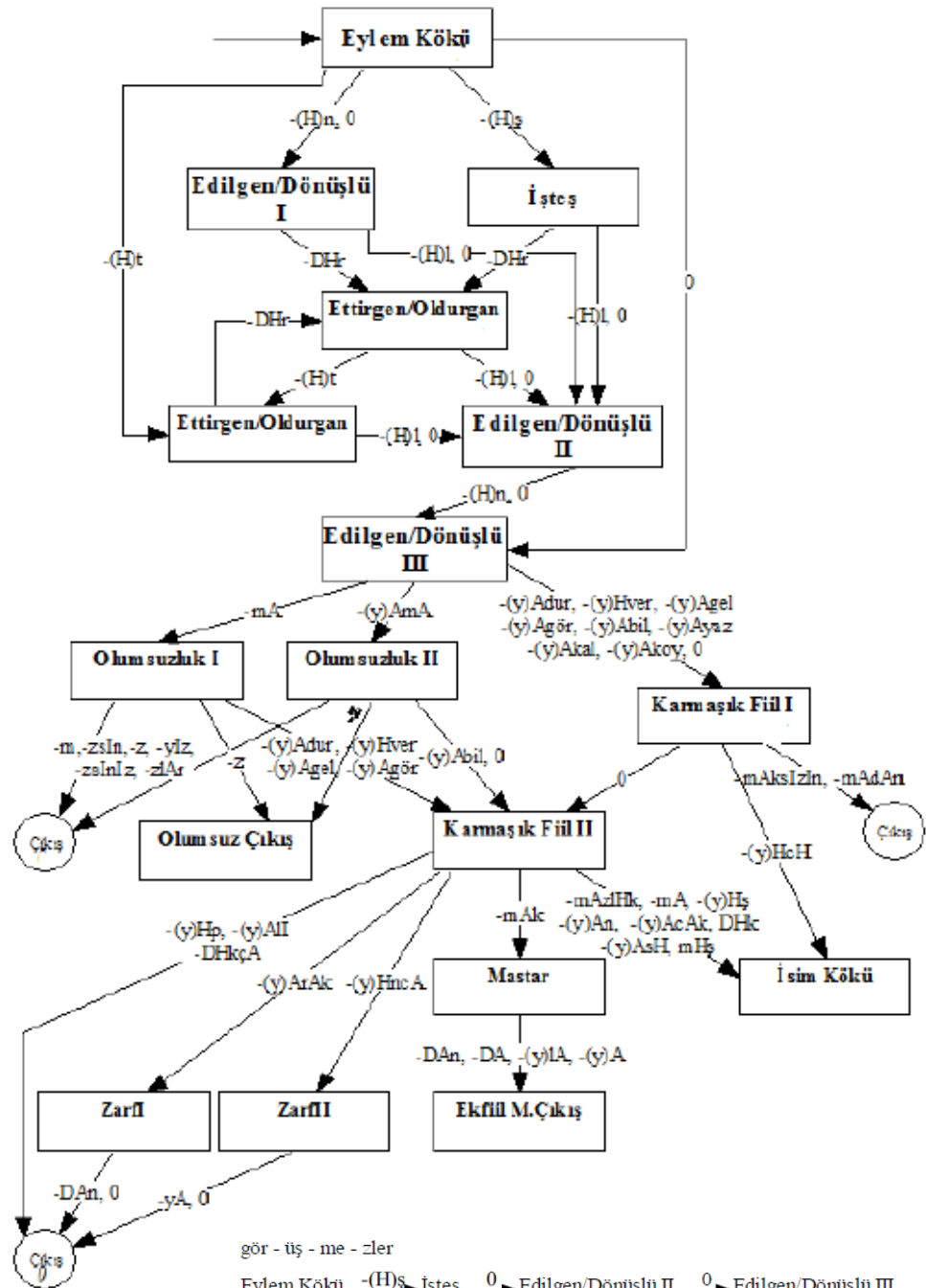
Fiil kökü-(H)yor → Zaman Tip II-(y)mHş → Bileşik Z.II-sHnHz → Çıkış

yapmaz - lar - sa

olumsuz -lAr → Bileşik Z.I-sA → Çıkış

Eylem Çekim Ekleri Kümesi

Ek No	Ek	Açıklama	Örnek
1	-m	1. tekil kişi	yap-ma-m
2	-zsln	2. tekil kişi	yap-ma-zsln
3	-z	3. tekil kişi	yap-ma-z
4	-ylz	1. çoğul kişi	yap-ma-yız
5	-zslnz	2. çoğul kişi	yap-ma-zsınız
6	-zlar	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-iver
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)Asl	sıfat fiil	oku-yası
29	-DHk	sıfat fiil	oku-duk
30	-mHş	sıfat fiil	oku-muş
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)IA	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	oldurğan	oku-t-tur-mak
44	-(H)t	ettirgen	oku-t-mak



oku - yucu - lar

Eylem Kökü \rightarrow Edilgen/Dönüşlü III \rightarrow Karmaşık Fiil I $-(y)HcH$ \rightarrow İsim Kökü

gör - üş - me - zler

Eylem Kökü $-(H)ş$ \rightarrow İşteş 0 \rightarrow Edilgen/Dönüşlü II 0 \rightarrow Edilgen/Dönüşlü III

$-mA$ \rightarrow Olumsuzluk I \rightarrow $-zlar$ \rightarrow Çıkış