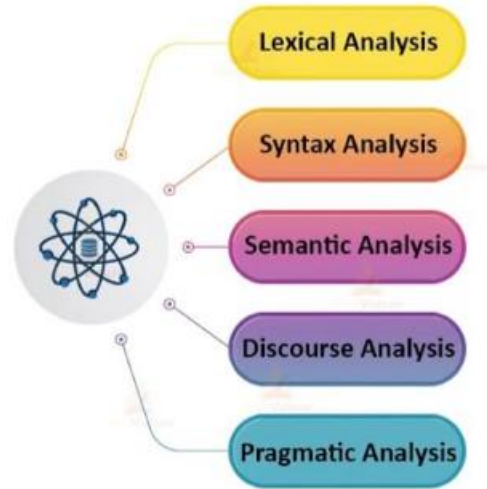


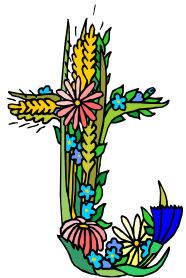
Doğal Dil İşleme Bilgi Düzeyleri-2



Prof.Dr. Banu Diri

POS

Sözdizimsel Analiz (Syntactic Analysis)



Sözdizimsel analiz, cümlenin yapısal tanımını oluşturabilmek için morfolojik analizin sonuçlarını kullanır.

Amacı, arka arkaya gelen kelimeler yığınının bu kelimeler yığınının ifade ettiği cümle birimlerini tanımlayan bir yapıya dönüştürmektir.

Cümle parçalarının doğru tespiti, anlam analizi için yararlı sonuçlar verir.

Grammer ve sözlükler, bilgisayardaki dilbilgisi modelini karakterize eder, algoritmik olarak işlenmesi de problemi matematiksel açıdan ifade eder.

-
- ✓ Bilgisayarla doğal dilin modellenmesinde anlamsal analizden önce kelimelerden oluşturulan yapının cümle olup olmadığı test edilmelidir.
 - ✓ Cümle oluşturmayan ifadelerden anlam çıkarmak imkansızdır. "*Kalem sarı koş kitap*" ifadesi bir anlam ifade etmez. Rastgele kelimelerin yan yana gelmesiyle cümle oluşturulamaz. Cümleleri oluşturan kelimelerin bir sonu olmalıdır.
 - ✓ Cümleler de sonlu yapılar olduğuna göre Sonlu Otomatalar ile gösterilebilirler. Bunun amacı, ardarda gelen kelimeler yığını bu kelimelerin ifade ettiği cümle öğelerini tanımlayan bir yapıya dönüştürmektir. Türkçe'deki cümleler için mümkün olan tüm sözdizimsel unsurlar dikkate alınmalıdır.

✓ Cümle unsurları, kelimeler, tamlamalar veya buna benzer cümle parçacıkları olabilir. Bu cümle parçacıklarının tespiti, anlamsal analiz için yararlı ipuçları vermektedir. Genellikle bir cümlede cümle parçacıkları ortak görev yaparlar.

Kelimelerin temel sözdizim fonksiyonları

İsimler(Nouns) dünyadaki varlıkları tanımlar (e.g. ‘people/insanlar’, ‘animals/hayvanlar’, ‘hat/şapka’).

Belirleyiciler(Determiners) ismi belirler (e.g. ‘the’, ‘a’, ‘an’).

Sıfatlar(Adjectives) isimleri niteleyen sözcükler olup isimlerle ortaya çıkar (e.g. ‘red’, ‘long’, ‘intelligent’). Tek başlarına kullanılmazlar. Örneğin “kırmızı”, bir renk ismidir, “iki”, bir sayı ismidir. Ancak bu sözcükler isimlerin özelliklerini bildirecek duruma gelirse sıfat olur. “kırmızı top”.

Füller(Verbs) bir oluşu, bir durumu kişiye bağlayarak anlatırlar. (e.g. ‘have/sahip olmak’, ‘threw/fırlatmak’, ‘walked/yürümek’).

Zarf/Belirteç(Adverbs) fiillerin niteliğini bildirir (e.g. ‘often’, ‘heavily’). “Güzel bir evde oturmak istiyorum.” cümlesinde “güzel” sözcüğü “ev” ismini nitelediği için sıfat, “Bu ev uzaktan daha güzel görünüyordu” cümlesinde ise “görünmek” fiilinin nasıl olduğunu bildirdiği için zarftır.

Edat/İlgeç(Preposition) Tek başına bir anlam taşımayan , ancak kendinden önceki sözcükle birlikte kullanıldığında belirli bir anlamı olan sözcüklerdir (e.g. ‘in’, ‘on’, ‘over’)(gibi, için, ile, kadar, karşı, göre, üzere, doğru, sanki, beri, sadece, ancak, yalnız, sonra,...)

Bağlaç(Conjunctions) Cümleleri veya aynı görevdeki sözcükleri birbirine bağlayarak aralarında anlam ilgisi kuran sözcüklere denir (e.g. ‘and’, ‘or’, ‘but’) (ile, ve, de, ama, fakat, ancak, yalnız, ki, hem...hem, ne...ne, çünkü, madem, ...).

Prepositions (edatlar)

of	540,085	through	14,964	worth	1,563	pace	12
in	331,235	after	13,670	toward	1,390	nigh	9
for	142,421	between	13,275	plus	750	re	4
to	125,691	under	9,525	till	686	mid	3
with	124,965	per	6,515	amongst	525	o'er	2
on	109,129	among	5,090	via	351	but	0
at	100,169	within	5,030	amid	222	ere	0
by	77,794	towards	4,700	underneath	164	less	0
from	74,843	above	3,056	versus	113	midst	0
about	38,428	near	2,026	amidst	67	o'	0
than	20,210	off	1,695	sans	20	thru	0
over	18,071	past	1,575	circa	14	vice	0

Pronouns (zamirler/adıllar)

it	199,920	how	13,137	yourself	2,437	no one	106
I	198,139	another	12,551	why	2,220	wherein	58
he	158,366	where	11,857	little	2,089	double	39
you	128,688	same	11,841	none	1,992	thine	30
his	99,820	something	11,754	nobody	1,684	summat	22
they	88,416	each	11,320	further	1,666	suchlike	18
this	84,927	both	10,930	everybody	1,474	fewest	15
that	82,603	last	10,816	ourselves	1,428	thyslf	14
she	73,966	every	9,788	mine	1,426	whomever	11
her	69,004	himself	9,113	somebody	1,322	whosoever	10
we	64,846	nothing	9,026	former	1,177	whomsoever	8
all	61,767	when	8,336	past	984	wherefore	6
which	61,399	one	7,423	plenty	940	whereat	5
their	51,922	much	7,237	either	848	whatsoever	4
what	50,116	anything	6,937	yours	826	whereon	2
my	46,791	next	6,047	neither	618	whoso	2
him	45,024	themselves	5,990	fewer	536	aught	1
me	43,071	most	5,115	hers	482	howsoever	1
who	42,881	itself	5,032	ours	458	thrice	1
them	42,099	myself	4,819	whoever	391	wheresoever	1
no	33,458	everything	4,662	least	386	you-all	1
some	32,863	several	4,306	twice	382	additional	0
other	29,391	less	4,278	theirs	303	anybody	0
your	28,923	herself	4,016	wherever	289	each other	0
its	27,783	whose	4,005	oneself	239	once	0
our	23,029	someone	3,755	thou	229	one another	0
these	22,697	certain	3,345	'un	227	overmuch	0
any	22,666	anyone	3,318	ye	192	such and such	0
more	21,873	whom	3,229	thy	191	whate'er	0
many	17,343	enough	3,197	whereby	176	whenever	0
such	16,880	half	3,065	thee	166	whereof	0
those	15,819	few	2,933	yourselves	148	whereto	0
own	15,741	everyone	2,812	latter	142	whereunto	0
us	15,724	whatever	2,571	whichever	121	whichsoever	0

Conjunctions (bağlaçlar)

and	514,946	yet	5,040	considering	174	forasmuch as	0
that	134,773	since	4,843	lest	131	however	0
but	96,889	where	3,952	albeit	104	immediately	0
or	76,563	nor	3,078	providing	96	in as far as	0
as	54,608	once	2,826	whereupon	85	in so far as	0
if	53,917	unless	2,205	seeing	63	inasmuch as	0
when	37,975	why	1,333	directly	26	insomuch as	0
because	23,626	now	1,290	ere	12	insomuch that	0
so	12,933	neither	1,120	notwithstanding	3	like	0
before	10,720	whenever	913	according as	0	neither nor	0
though	10,329	whereas	867	as if	0	now that	0
than	9,511	except	864	as long as	0	only	0
while	8,144	till	686	as though	0	provided that	0
after	7,042	provided	594	both and	0	providing that	0
whether	5,978	whilst	351	but that	0	seeing as	0
for	5,935	suppose	281	but then	0	seeing as how	0
although	5,424	cos	188	but then again	0	seeing that	0
until	5,072	supposing	185	either or	0	without	0

Auxiliaries (yardımcı fiiller)

can	70,930	might	5,580	shouldn't	858
will	69,206	couldn't	4,265	mustn't	332
may	25,802	shall	4,118	'll	175
would	18,448	wouldn't	3,548	needn't	148
should	17,760	won't	3,100	mightn't	68
must	16,520	'd	2,299	oughtn't	44
need	9,955	ought	1,845	mayn't	3
can't	6,375	will	862	dare	??
have	???				

POS (Part of Speech)

Doküman içerisinde yer alan her bir kelime **POS** olarak adlandırılır. Her POS kendisine verilen **TAG**'ler (etiket) ile ifade edilir ve sözdizimsel yapı çıkarılırken kullanılır.

Benzer sözdizimsel davranışlar ile kelimelerin sınıflandırılması **syntactic** veya **grammatical categories** veya **Parts of Speech (POS)** olarak adlandırılır.

İngilizce cümle öğelerinin etiketlenmesi için hazırlanmış olan ***Penn Treebank*** ve ***Brown*** derlemi kullanılır.

Brown TAG kümesi

Tag	Description	Example	Tag	Description	Example
CC	Coordin. Conjunction	<i>and, but, or</i>	SYM	Symbol	<i>+, %, &</i>
CD	Cardinal number	<i>one, two, three</i>	TO	"to"	<i>to</i>
DT	Determiner	<i>a, the</i>	UH	Interjection	<i>ah, oops</i>
EX	Existential 'there'	<i>there</i>	VB	Verb, base form	<i>eat</i>
FW	Foreign word	<i>mea culpa</i>	VBD	Verb, past tense	<i>ate</i>
IN	Preposition/sub-conj	<i>of, in, by</i>	VBG	Verb, gerund	<i>eating</i>
JJ	Adjective	<i>yellow</i>	VBN	Verb, past participle	<i>eaten</i>
JJR	Adj., comparative	<i>btgger</i>	VBP	Verb, non-3sg pres	<i>eat</i>
JJS	Adj., superlative	<i>wildest</i>	VBZ	Verb, 3sg pres	<i>eats</i>
LS	List item marker	<i>1, 2, One</i>	WDT	Wh-determiner	<i>whitch, that</i>
MD	Modal	<i>can, should</i>	WP	Wh-pronoun	<i>what, who</i>
NN	Noun, sing. or mass	<i>llama</i>	WPS	Possessive wh-	<i>whose</i>
NNS	Noun, plural	<i>llamas</i>	WRB	Wh-adverb	<i>how, where</i>
NNP	Proper noun, singular	<i>IBM</i>	\$	Dollar sign	<i>\$</i>
NNPS	Proper noun, plural	<i>Carolinas</i>	#	Pound sign	<i>#</i>
PDT	Predeterminer	<i>all, both</i>	"	Left quote	<i>(' or ")</i>
POS	Possessive ending	<i>'s</i>	"	Right quote	<i>(' or ")</i>
PRP	Personal pronoun	<i>I, you, he</i>	(Left parenthesis	<i>([{ <</i>
PRP\$	Possessive pronoun	<i>your, one's</i>)	Right parenthesis	<i>(] , } >)</i>
RB	Adverb	<i>qutckly, never</i>	,	Comma	<i>,</i>
RBR	Adverb, comparative	<i>faster</i>	.	Sentence-final punc	<i>(. ! ?)</i>
RBS	Adverb, superlative	<i>fastest</i>	:	Mid-sentence punc	<i>(: ; ... - -)</i>
RP	Particle	<i>up, off</i>			

Türk Dilleri için TAG kümesi

	Türk	Tatar	Mıgar	Başkir	Başkir	Kazak	Leipzig
3. tekil kişi	A3sg	SG	SG	SG	Şg	N0	SG
3. çoğul kişi	A3pl	A3pl	PL-JIÄp	PL	PI	N1	PL
İyelik							
1. tekil iyelik eki	P1sg	POSS_1-SG	1SG	POSS.1SG	Poss1 Sg P1	S1	POS, POSS
2. tekil iyelik	P2sg	POSS_2-SG	2SG	POSS.2SG	Poss1 Sg P2	S2	
3. tekil iyelik	P3sg	POSS_3-SG	3SG	POSS.2SG	Poss1 Şg P2	S3	
1. çoğul iyelik	P1pl	POSS_1-PL	1PL	POSS.1PL	Poss1 Pl P1	S5	
2. çoğul iyelik	P2pl	POSS_2-PL	2PL	POSS.2PL	Poss1 Pl P1	S6	
3. çoğul iyelik	P3pl	POSS_3-PL	3PL			S7	
2. tekil iyelik resmi						S4	
2. çoğul iyelik resmi						S8	
Durum Ekleri							
Yalın	Nom	NOM		NOM	Nom	C1	NOM
İlgi	Gen	GEN	GEN	GEN	Gen	C2	GEN
Yönelme	Dat	DIR	DAT	DAT	Dat	C3	DAT, DIR
Sınırlı yönelme		DIR_LIM					
Belirtme	Acc	ACC	ACC	ACC	Acc	C4	ACC
Ayrılma	Abl	ABL	ABL	ABL	Abl	C6	ABL
Bulunma	Loc	LOC	LOC	LOC	loc	C5	LOC
Enstrüman	Ins					C7	INS
Yön							DIR
Eşitlik	Equ		ADV				
Benzerlik							SIM
Birliktelik							COM
Nitelikler							
sız	Minor-pos Without	ATTR_ABES	ABES	ABE	Abs		ABS DEL
ile	With	ATTR_MUN	ATR				ATT, ATTR
daki	Rel	ATTR_LOC	ATR		Loc2-6		
aitlik		ATTR_GEN		IPOSS		S9	

			Derivation				
Soyut ad	Noun	NMLZ	NOMIN	DERIV. ABSTR	Abst		NMZ, NZ, NOMI
İşi yapan	Agt	PROF		DERIV. AGE NS			AG, AGT
Küçültme eki	Dim AFC	DIM			Dimhp		DIM
Karşılaştırmanın Derecesi							
Karşılaştırma	Adj, Adv	COMP	COMP	COMP	Com	SES	COMP
karşılaştırma belirteç						USS	
Üstünleme önad	Adj					SEA	SUP(L)
Üstünleme belirteç						USA	
Sayı Türleri							
Topluluk sayı önadları	Col	NUM_COLL		COL		SNZ	COL
Olağan sayılar	Num+Ord	NUM_ORD	ORD	ORD		SNR	
Üleştirme sayıları	Num+Dist	NUM_DISR		DIVIS			DIST, DISTR
Kesirli sayılar		NUM_APPR		APRX			APRX, APPR
Ayrıştırma sayıları				DIST			
Kişiler (ad ve eylem)							
1. tekil kişi	A1sg	1SG	1.SG	1SG	P1	P1	1, 1s
2. tekil kişi	A2sg	2SG	2.SG	2SG	P2	P2	2, 2s
3. tekil kişi	A2sg	2SG	3.SG	3SG	P3	P3	3, 3s
1. çoğul kişi	A1pl	1PL	1.PL	1PL	P5	P5	1p
2. çoğul kişi	A2pl	2PL	2.PL	2PL	P2	P6	2p
3. çoğul kişi	A3pl	3PL	3.PL	3PL	P3	P7	3p
2. tekil kişi kibar						P4	POL
2. çoğul kişi kibar						P8	
ek eylem geniş zaman				PRED	Pred		PRED

Soru							
Soru	Post+Ques	INT	Q	INTERROG	Q		INT, Q
Merak		INT_MI					MIR
Olası	Able	PROB		UNCERT			PROB, INDF
Gibi	Adj	SIM		ASSIM	Comp1		SIM
-sı, -si	Adv	SIM_2					
Benzer		SIM_3					
-trak	Adj						
Zamanlar							
şimdiki	Pres	PRES	IPFV	PRES	Prs	T1	PRES, PRS
Geniş							
-di'li geçmiş	Past	PST_DEF	PRET	PST_DEF	Pstdef	T3	DEF
-miş'li geçmiş	Narr	PST_INDF	PFCT	PST_INDF	Pstindf	T3	INDF
Gelecek z (kesin)	Fut	FUT_DEF	POT2	FUT_DEF	Futdef		FUT
Gelecek z (belirsiz)		FUT_INDF	POT1	FUT_INDF	Futindf	T2	
Olumsuz gelecek		FUT_INDF_NEG					
Geniş z (geçişli)						T2(-A)	TR, TRANS
Geçmiş z (geçişli)						T3	
Olumsuz							
Olumsuz		NEG	NEG	NEG			
Olumsuz değil							
			Participles and other attributive verb				

İngilizce ile Türkçe arasındaki farklılıklar

- İngilizcede tanımlık (the, a) vardır, Türkçede yoktur. Belirsiz *bir* vardır.
- İngilizcede ön durum (preposition) (of, in, by) vardır, Türkçede ön durum yoktur.
- İngilizcede karşılaştırma ve üstünleme ile önad oluşturma biçimi Türkçeden farklıdır. Örneğin *big, bigger, biggest* - büyük, daha büyük, en büyük
- İngilizcede karşılaştırma ve üstünleme belirteçlerini oluşturma biçimi Türkçeden farklıdır. Örneğin *fast, faster, fastest* - hızlı, daha hızlı, çok hızlı
- İngilizcede parçacık (up, off) vardır, Türkçede parçacık yoktur.
- Türkçede tüm eylemler kurallıdır; eylem zaman ve kişi bilgisini içerir. Bu nedenle her bir durum için özel kısaltma gerektirir. İngilizcede eylemler kurallı ve kuralsız olarak iki sınıfa ayrılır. Kurallı olanlarda yalnız üçüncü tekil kipte kişi bilgisi ve zaman bilgisi bulunur.
- İngilizcedeki tümcecikleri birbirine bağlayan *which, that, what, who, whose, how, where* gibi yan tümceyi ana tümceye bağlayan ögeler Türkçede yoktur.
- Türkçe eklemeli bir dildir, özellikle yapım ekleri çoktur. İngilizcedeki yapım eklerinin sayısı görece azdır.

Sözcük Öbek yapısı-Sözcük dizimi-TR (Phrase structure)

- Kelimeler öbek içerisinde belli bir sırada yer alır.
- Türkçe’de cümleler en genel şekliyle **özne**, **nesne** ve **yüklem** bileşenlerinden oluşur. Cümledeki öğelerin yerini değiştirebiliriz (özgür sıralı söz dizimi).
- Cümleye eklenmek istenen anlamlar arttıkça cümleler, özne, yer tamlayıcısı, **zarf** tamlayıcısı, nesne ve yüklem gibi bileşenleri de içerir.
- Cümlelerin anlamını kuvvetlendiren cümle dışı bileşenler de (**bağlaç**, **edat**, vb) cümlede bulunabilir.
- Türkçe’de özne ile yüklem cümlelerin 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.

Bu TAG'lerden oluşabilecek cümleler; öğelerinin dizilişlerine göre bir çok varyasyonlar oluşturabilirler. Bu varyasyonlar;

❖ Ö-D-Z-N-Y, Ö-D-N-Z-Y, Ö-Z-D-N-Y, Ö-Z-N-D-Y, Ö-N-Z-D-Y, Ö-N-D-Z-Y olabilir.

❖ Bunlara tüm öğeleri içermeyen iki, üç veya dört öğeden oluşan cümleler de eklenebilir.

İki unsur için: Ö-Y

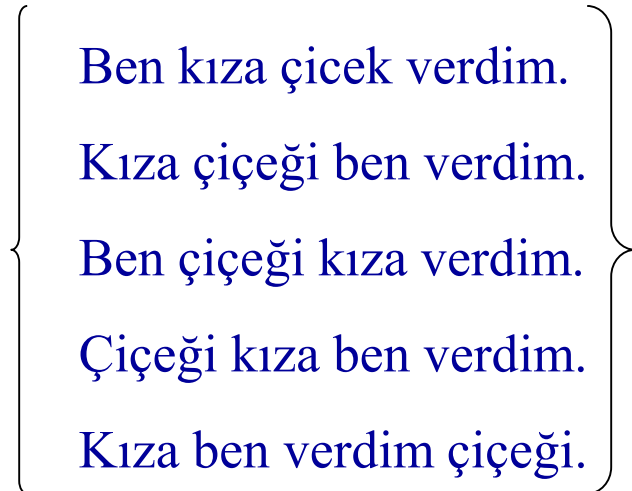
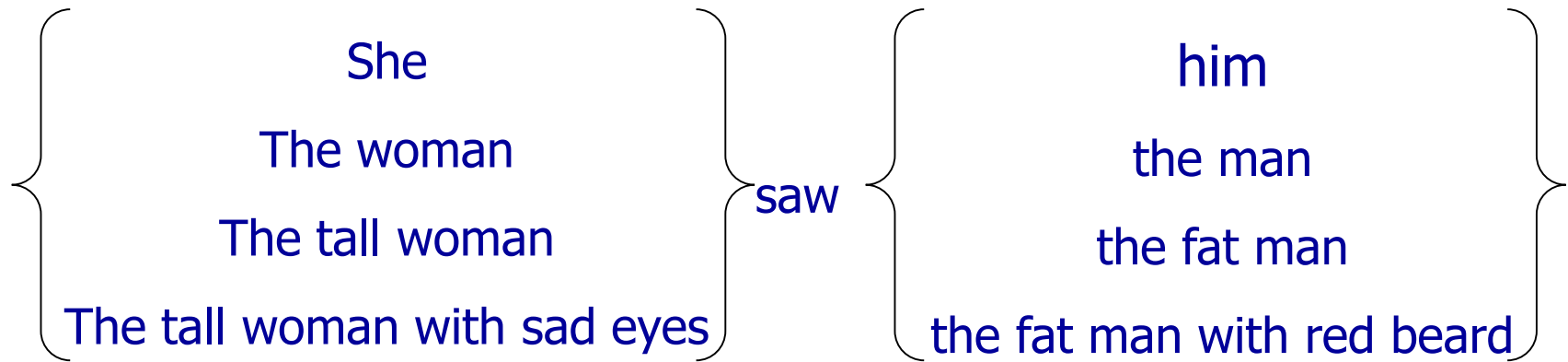
Üç unsur için Ö-N-Y, Ö-D-Y, Ö-Z-Y

Dört unsur için Ö-N-D-Y, Ö-N-Z-Y, Ö-D-N-Y, Ö-D-Z-Y, Ö-Z-N-Y, Ö-Z-D-Y

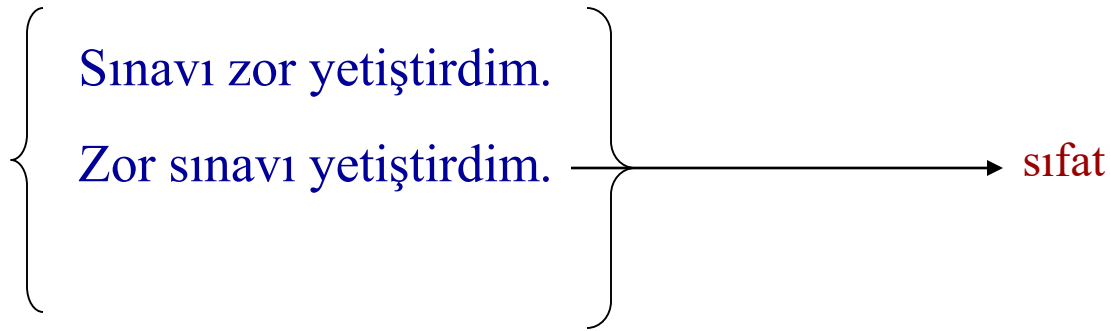
❖ Bu varyasyonların hepsi $\{\ddot{O}\lambda\lambda\lambda Y\}$ şeklinde gösterilebilir.

Öbek yapısı-Sözcük dizimi (Phrase Structure)

- İngilizce’de sözcük sırası (word order) son derece katıdır.



Serbestliğin kısıtlı olduğu durumlar mevcuttur. Belirteç'lerin (zarf) kullanımı gibi.



Problem ??

Kelimeler birden fazla sınıfa dahil olabili: ***this***

This is a nice day = PRP (*personal pronoun*)

This day is nice = DT (*determiner*)

You can go *this* far = RB (*adverb*)

Sözcük Öbekleri

- Tamlamalar
 - Ad Tamlamaları
 - Önad Tamlamaları
- Yüklemin Türüne Göre Tümceler (Eylem ve Ad Tümceleri)
- Öğelerin Dizilişine Göre Tümceler

%46 ÖNY, %7 NÖY, %20 NYÖ, %10 YNÖ
- Tümcelerın Anlamsal Sınıfları (Olumlu, Olumsuz, Bildirme, Soru, Emir, İstek, Ünlem cümleleri)
- Birleşik Tümceler

Tümce Çözümü

- ❑ Bağlam Bağımsız Dilbilgisi (Context-Free Grammer) *
- ❑ Sözcesel İşlevsel Dil Bilgisi
- ❑ Ulamsal Dil Bilgisi
- ❑ Başsözcüğe Dayalı Öbek Yapısal Dilbilgisi (Head-Driven Phrase Structure Grammer)
- ❑ Bağlı Dil Bilgisi

Bağlam Bağımsız Dilbilgisi (Context-Free Grammar)

- Bir kelime ile cümlelerin söz dizimini betimleyen ağaç yapısı arasındaki bağıntıyı belirler
- Ağaç Yapılı Derlemi kullanır (cümle ile buna karşılık gelen ağaç yapısını barındırır)
- Kullanılan Çözümler
 - Kural Tabanlı Çözümler (yukarıdan aşağıya, aşağıdan yukarıya ayrıştırma)
 - Olasılık Tabanlı Çözümler (her tümce ve ağaç için bir olasılık değeri atanır)

Bir cümle/tümce temel olarak

özne, nesne, tümleç ve yüklemden oluşur.

Türkçe

Türkçe
Fince

Aynı dil ailesi
Ural Altay

olarak görür

Kereha pendakewo karenajuu tosite miru
O kalemini tek silahı olarak görür

Korece

그는 그의 펜을 그의 유일한 팔로 생각한다.
O mini kalem i(ını) tek silah olarak görür

Fince

ÖNTY
ÖYNT

Han pitaa kynaansa ainoana aseenaan
O görür kalemini olarak tek silahı

Çince

他把自己的笔看作成他唯一的武器
O mini onun kalem görür olarak onun tek silah

İngilizce
Çince
ÖYNT

İngilizce

He regards his pen as his only arm
O görür kalemini olarak tek silahı

Türkçe Cümlelerin Öğelerinin Bulunması

Türkçenin Cümle Biçimi

Türkçe, dört farklı cümle/tümce sınıfına sahiptir.

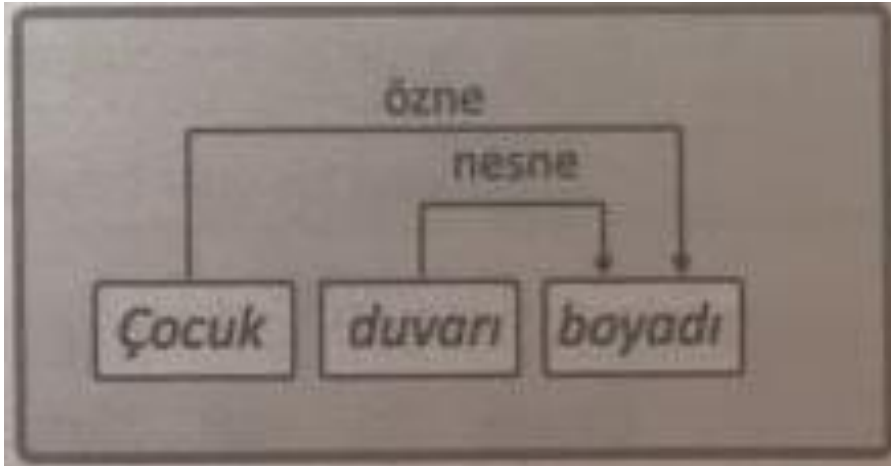
$$T = \{YT_1, YT_2, \dots, YT_n\} TT$$

YT: Yan Tümce

TT: Temel Tümce

T: Tümce

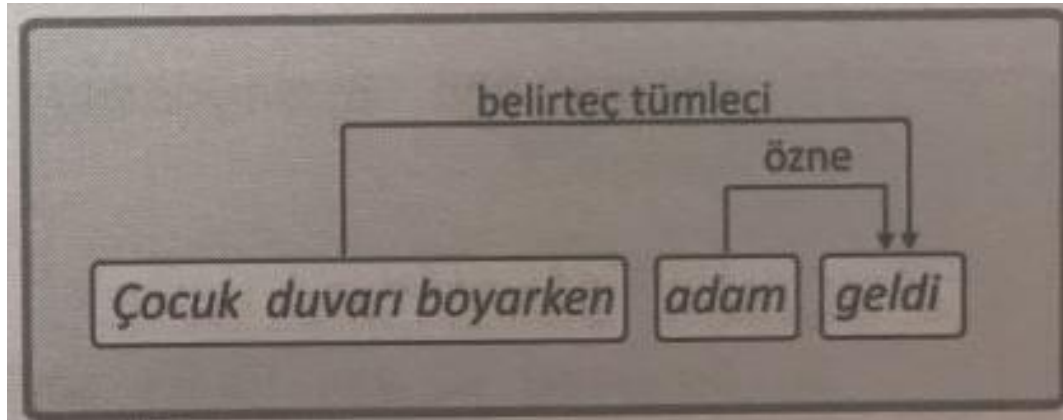
Basit Tümce



$T = TT$

Çocuk duvarı boyadı

Ortak Ögesiz Birleşik Tümce

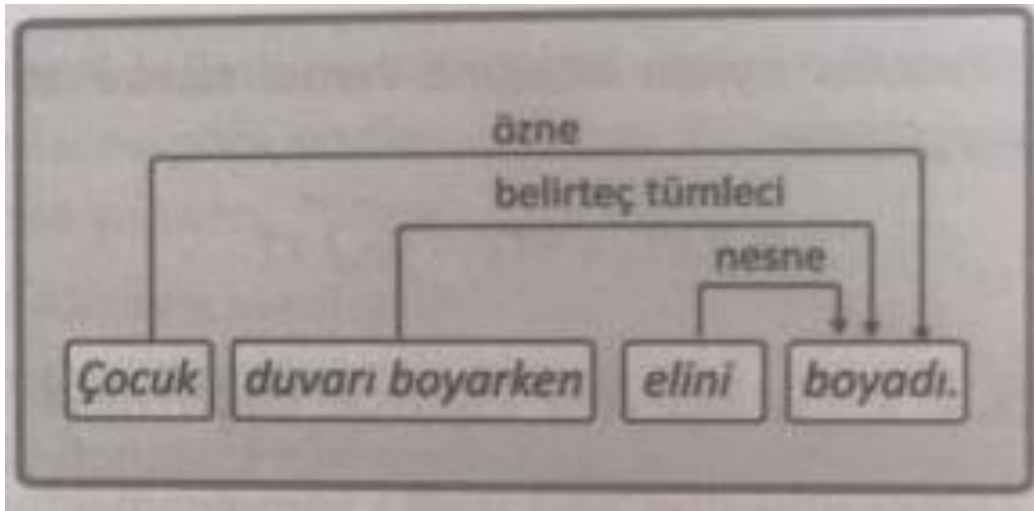


$$T = \{ YT_1 + YT_2 + \dots + YA_N \} TT$$

{Çocuk duvarı boyarken} {adam geldi.}

$$T = \{ YT_1 + YT_2 \} TT$$

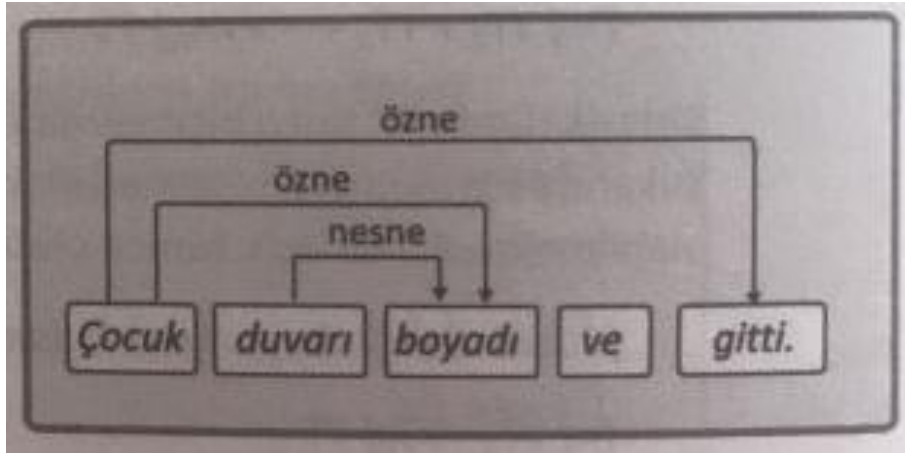
Ortak Ögeli Birleşik Tümce



$$T = \text{ÖZN} \{ YT_1 + YT_2 \} TT$$

{Çocuk duvarı boyarken} {elini boyadı.}

Ortak Özneli Bağımlı Sıralı Tümce

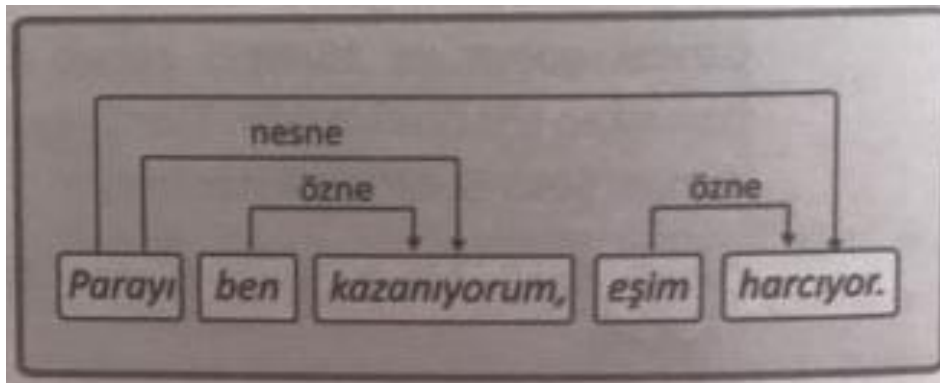


$$T = TT_1 + TT_2 + \dots + TT_n$$

{Çocuk duvarı boyadı} ve {gitti.}

$$T = \text{ÖZN}\{TT_1 + TT_2\}$$

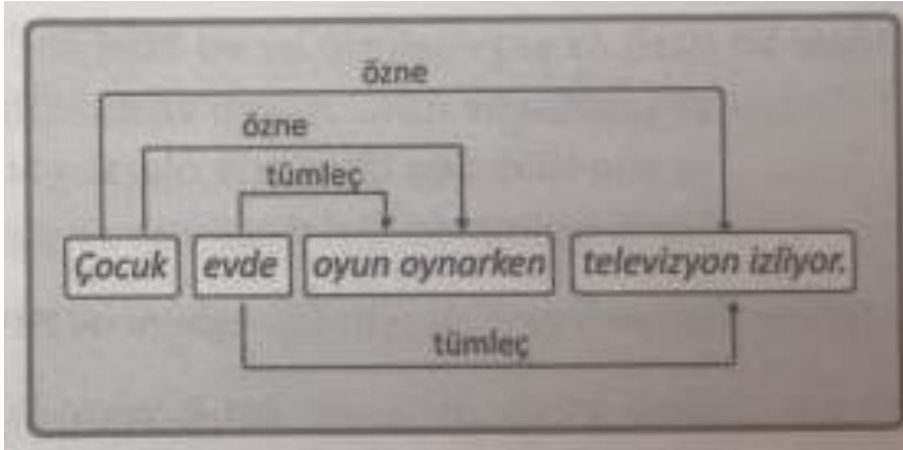
Ortak Nesneli Bağımlı Sıralı Tümce



{Parayı ben kazanıyorum,} {eşim harcıyor.}

$$T = \text{NES}\{TT_1 + TT_2\}$$

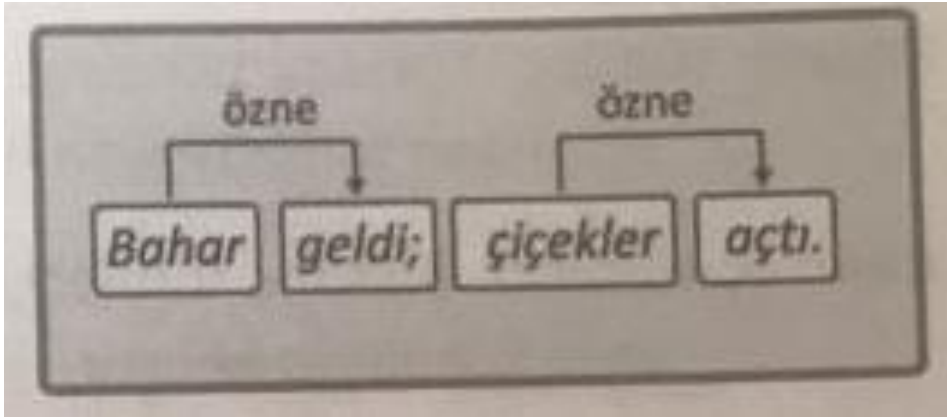
Tümceler arasındaki bağı kuran birden fazla öge olabilir



{Çocuk evde oyun oynarken} {televizyon izliyor.}

$$T=NES\{TT_1+TT_2\}$$

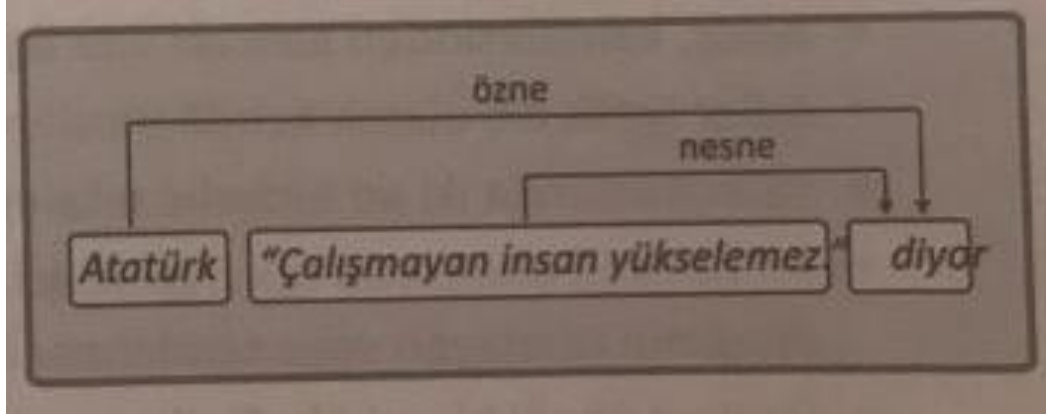
İki tümceyi birbirine bağlayan
öğeler özne ve tümleçtir.



{Bahar geldi,} {çiçekler açtı.}

Sıralı tümceler birbirinden de
bağımsız olabilir.

Girişik Tmce



$$T = NES\{YT_1\} TT_1 + TT_2 + \dots + TT_n$$

{Atatrk {"Çalışmayan insan ykselemez."} diyor}

Tümcenin Öğeleri

- Tümcelerın büyük çoğunluğu yan tümcelerden oluşmaktadır.
- Yan Tümce, özne ve yüklemden oluşur.
- Bir tümce bir veya daha fazla yan cümleden oluşmaktadır.
- Bir tümce, iki kurucudan oluşur, Özne ve Yüklem.

Özne: Bir basit özne genellikle bir ad öbeğidir. Türkçe tümcelerde özne genellikle ad öbeğidir (AÖ) ve yüklem ile özne uyumu vardır. Özne genellikle adın yalın durumunda olur. Türkçe tümce ve yan tümcede özne açık olarak görünmeyebilir, eylem içinde belirtilebilir. Özne sözcük öbeği biçiminde olabilir.

Yüklem: Öznenin gerçekleştirdiği eylemi belirtir.

Nesne: Tümce içinde öznenin yaptığı veya yüklem tarafından bildirilen iş veya oluştan etkilenen öğelerdir. Nesneler, belirtili ve belirtisiz nesne olarak ikiye ayrılırlar. Belirtili nesneler tümce içinde -i durum ekini alırken belirtisiz nesneler yalın durumdadır.

Dolaylı tümleş: Yüklemi yönelme (-e durumu), bulunma (-de durumu) ve ayrılma (-den durumu) açısından tamamlayan öğelerdir.

Belirteç tümleci: Belirteç tümleci, yüklemın belirttiği eylemin zaman, durum, derece, nicelik, yön veya koşulunu belirtir.

Tümcenin öğelerini ÖYD bağlamında bulmaya ad öbeklerinden başlayabiliriz. Türkçede ad öbekleri, ad (cins ve özel adlar), adıl, belirteç veya adı niteleyen, belirten ya da tümleyen sözcük öbeği biçiminde olabilir. Türkçe AÖ'lerin bulunmasında kullanılan kurallar aşağıda sıralanmıştır:

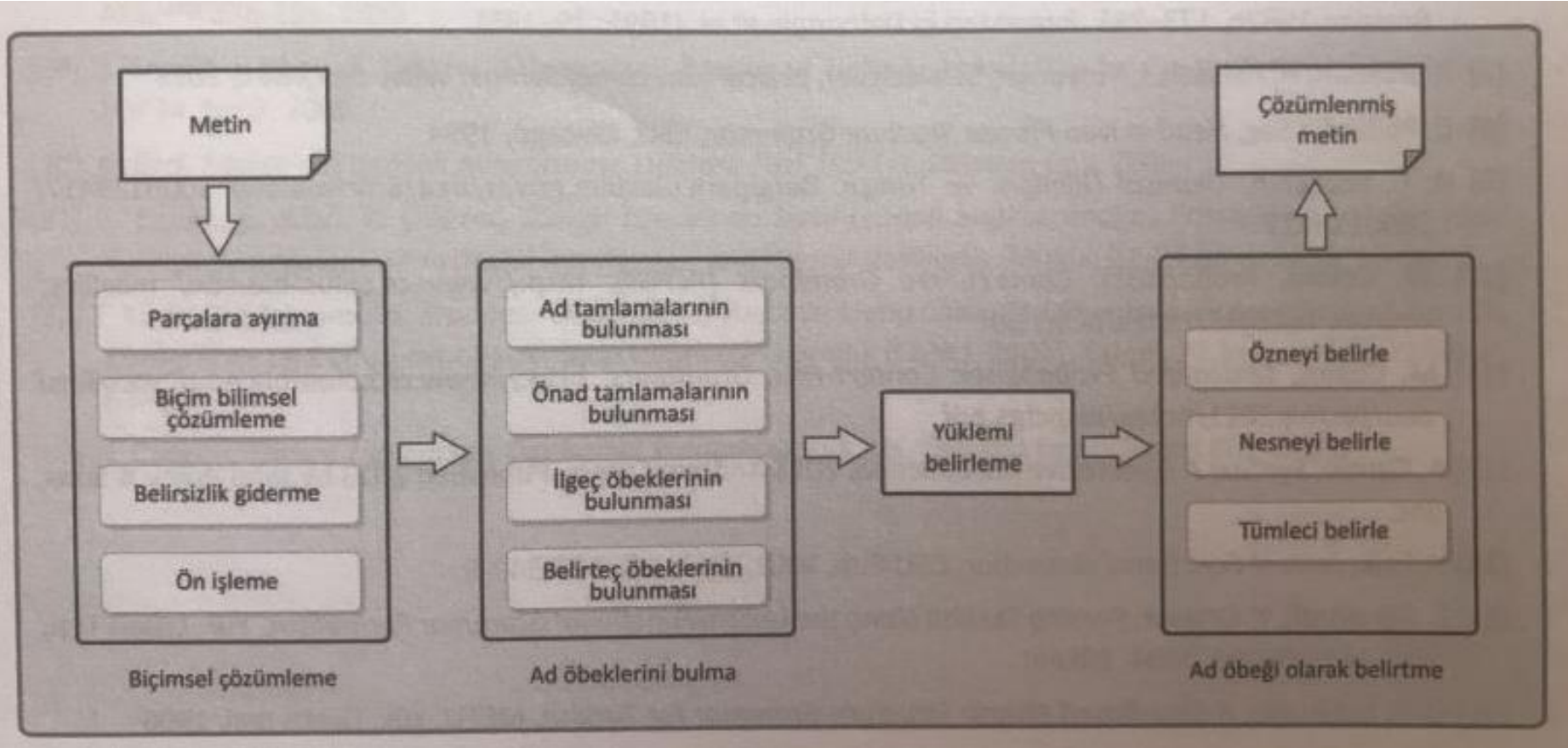
- Ad öbeği tek bir addan, adıl veya belirteçten oluşabilir: *annem*.
- Önad, adı nitelemek üzere adın önünde yer alır: *büyük kırmızı ev*
- Adlar kendilerinden önce gelen adılar ile ad kümeleri oluşturur: *benim elim*
- Adlar, kendilerinden sonraki ada ilgi eki ile bağlanabilir: *çocuğ+un kitabı*
- Adlar iyelik eki alarak kendilerinden önce gelen ada bağlanabilir: *okul zil+i*
- Yalın durumda iki ad birbirini niteleyebilir: *taş ev*
- Ad öbeğini oluşturan tek başına ad soylu sözcükler olabileceği gibi ad öbekleri de bir ad öbeğinin tamlayan veya tamlananı olabilirler: *çocuğun oyuncağının tekerleği*
- Özel ad öbeği bir ad öbeği oluşturabilir: *Millî Eğitim Bakanlığı Genel Sekreteri*
- Ad öbekleri bağlaçlarla birbirine bağlanmış ad öbeklerinden, tamlayan veya tamlanandan oluşabilir: *sarı veya mavi kalem*
- Tümcede yer alan belirteçler ad öbekleri oluştururlar: *daha büyük bir ev*
- Zaman bildiren öbekler ad öbeği olarak değerlendirilir: *19 Mayıs 1919 saat 15:30'da*

- Ad öbeğine eklenen ilgeçler ad öbeğinin içinde kabul edilir: *büyük bir sevinçle*
- Türkçede eylemsiler tümce içerisinde ad, önad veya belirteç gibi görevlerde bulunurlar.

Tümcenin öğeleri içinde yüklemnin ağırlığı bulunmaktadır. Bu nedenle ilk olarak yüklemnin bulunması yeğlenir. Yüklemini bulmak üzere üç kural oluşturulmuştur:

- 1) Tümcede çekimli eylemler yüklem görevindedir.
- 2) Ek-eylem eki alan ad kümeleri yüklem görevindedir.
- 3) Birleşik eylemleri bulmak için TDK'nin birleşik eylem, deyim ve atasözleri sözlüğündeki birleşik eylemlerden faydalanılabilir.

AÖ: Ad Öbeği
ÖYD :
Bağlamdan
Bağımsız
Dilbilgisi



Örnek “Bir cümle isim veya fiil grubundan oluşur.”, “isim grubu isim ve/veya sıfat, fiil grubu da isim grubu ve/veya fiil den oluşur”.

Bu gramerde “Küçük çocuk kırmızı top aldı” cümlesinin çözümü

$C \rightarrow \dot{I}G \ FG$: (cümle isim ve fiil grubundan oluşur)

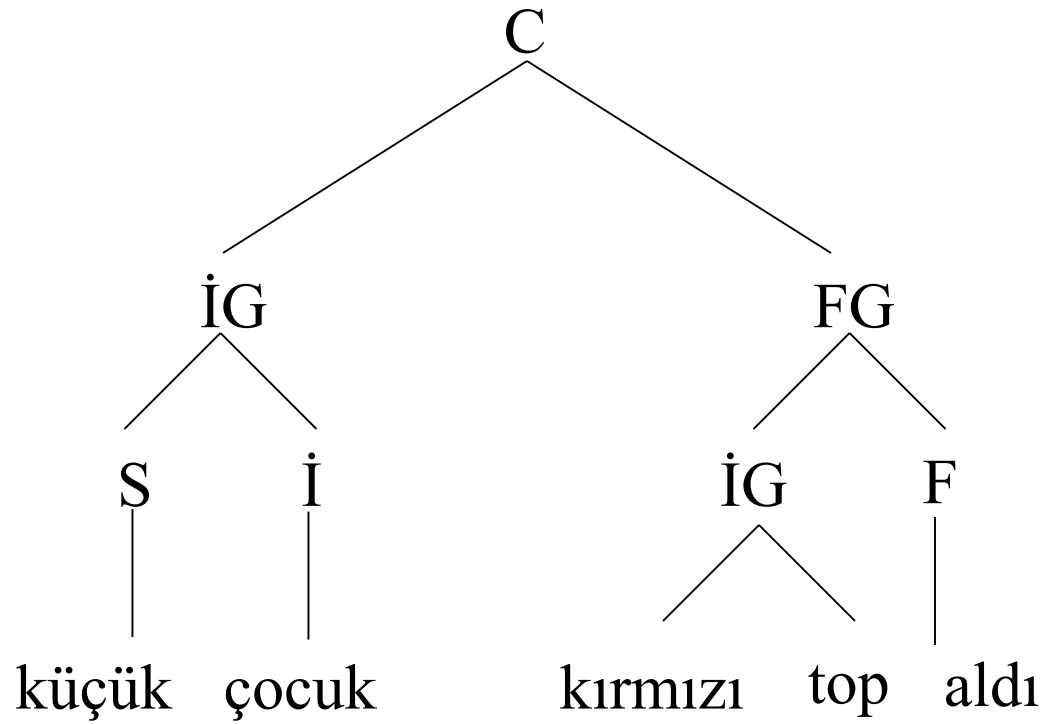
$\dot{I}G \rightarrow S \ \dot{I}$: (isim grubu \rightarrow sıfat+isim)

$FG \rightarrow \dot{I}G \ F$: (fiil grubu \rightarrow isim grubu + fiil)

$S \rightarrow$ küçük|kırmızı

$\dot{I} \rightarrow$ çocuk|top

$F \rightarrow$ aldı



Brown Tags

NN – singular noun

NNP – proper nouns

NNS – plural nouns

NR – adverbial nouns (e.g. 'home')

JJ - adjective

AT – articles

VB – verb, base form

VBD – verb third person singular (e.g. 'likes')

RB – adverbs

IN - preposition

- S --> NP VP
- NP --> AT NNS |
AT NN |
NP PP
- VP --> VP PP |
VBD |
VBD NP
- PP --> IN NP

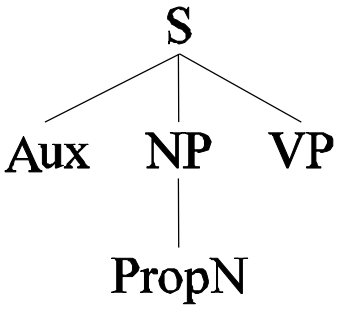
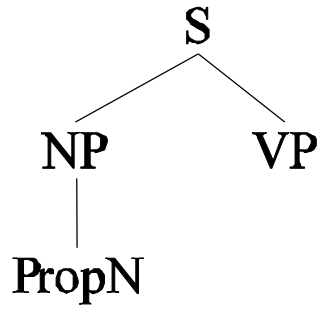
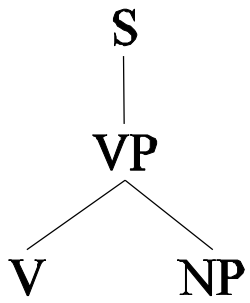
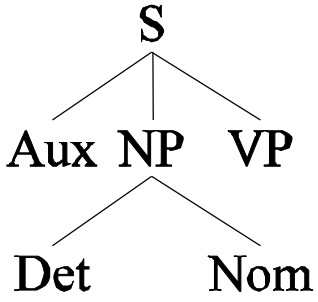
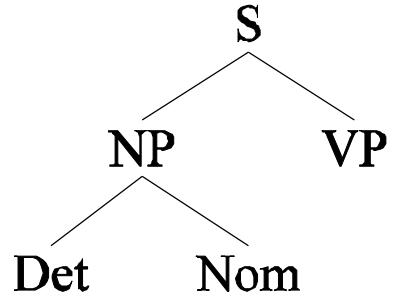
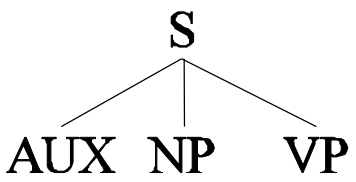
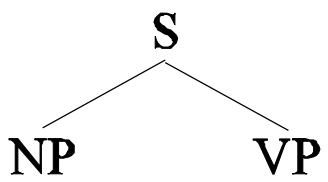
The Grammar

- AT --> the
- NNS --> children |
students |
mountains
- VBD --> slept |
ate |
saw
- IN --> in |
of
- NN --> cake

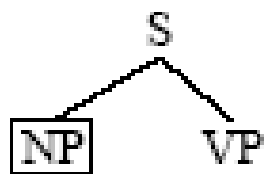
The Lexicon

Bir cümle için :Yukarıdan-Aşağıya Ayırıştırma (Top-Down Parsing, Left-to-Right, Depth-First)

S

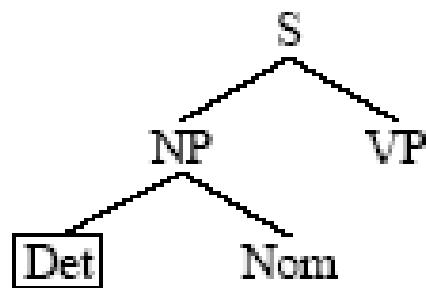


[S]

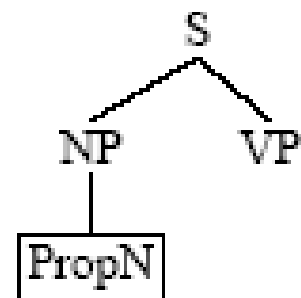


[Does]

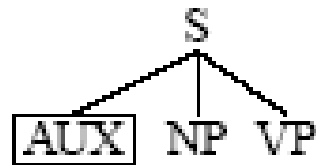
[Does]



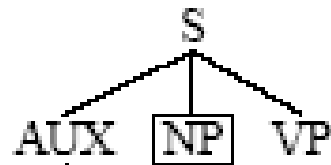
[Does]



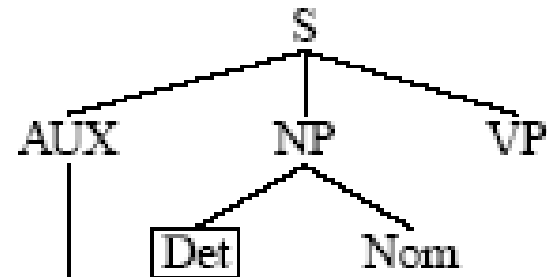
[Does]



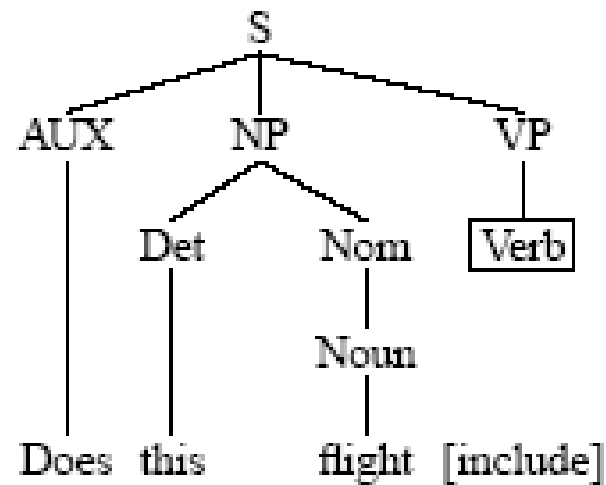
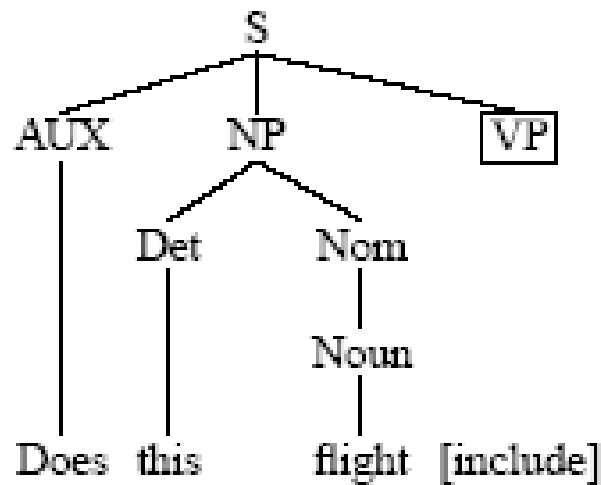
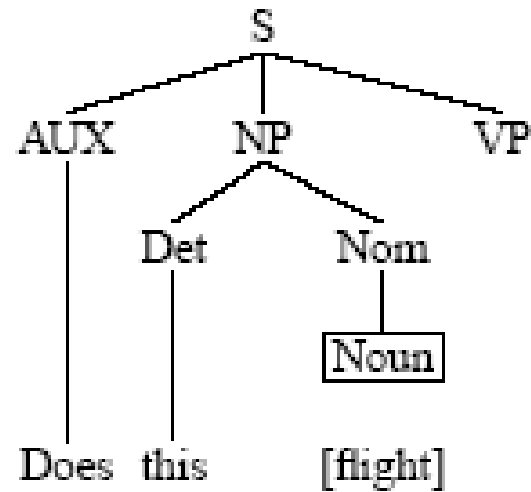
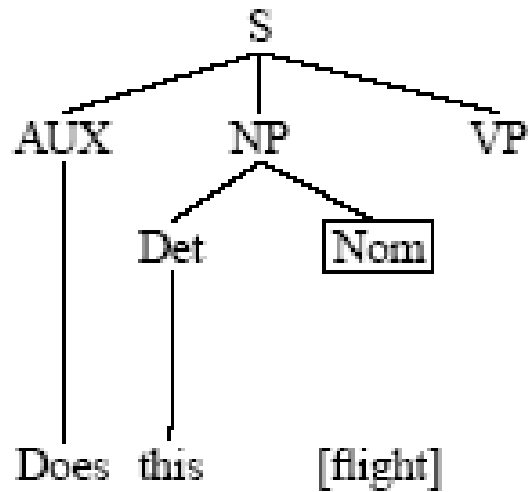
[Does]

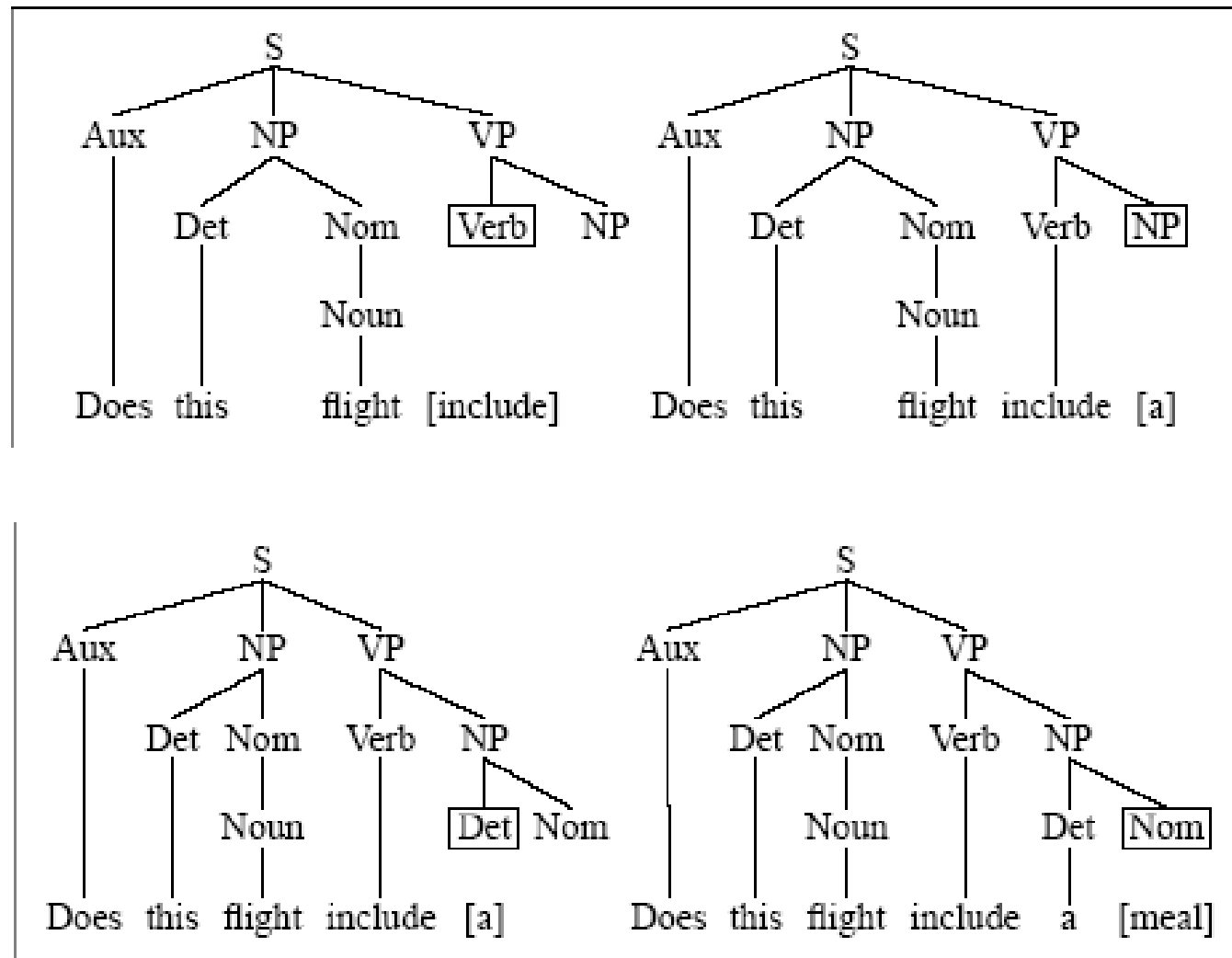


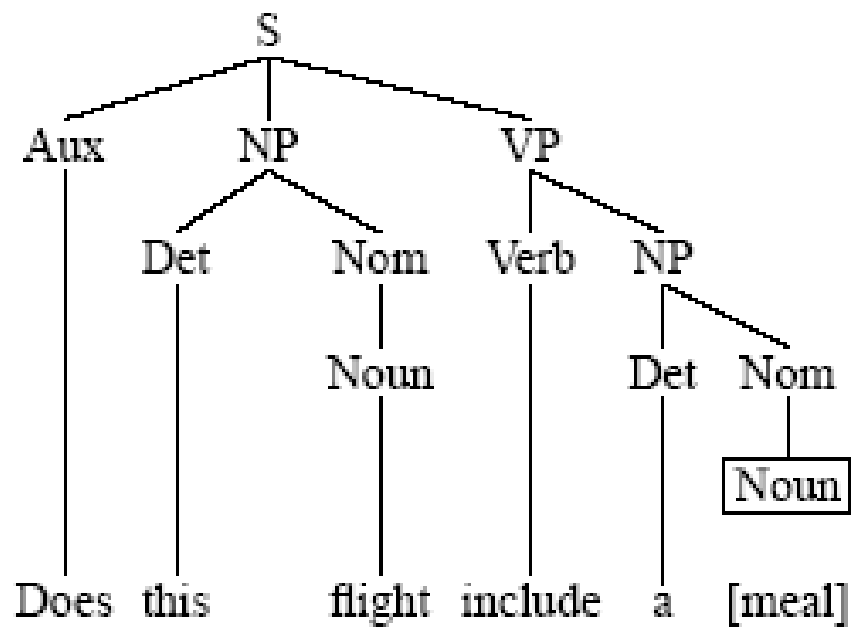
Does [this]

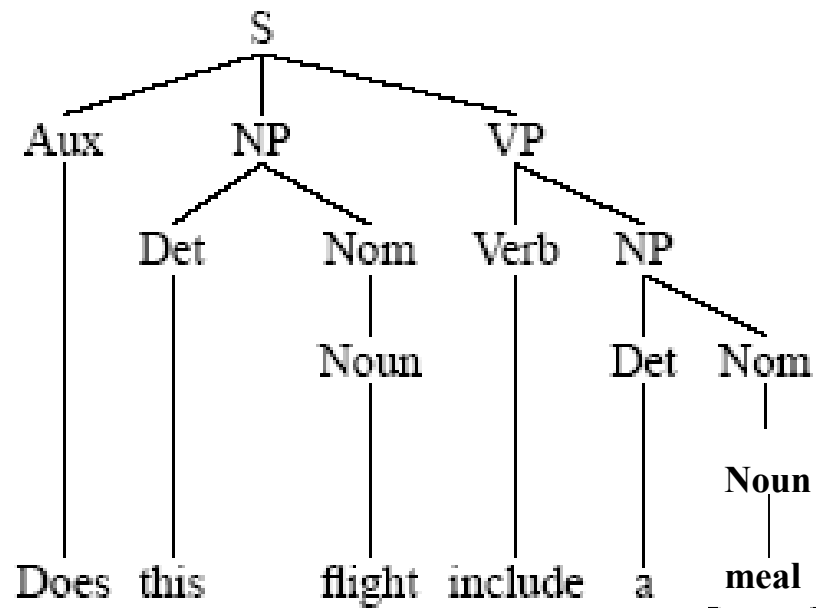


Does [this]









S → Aux NP VP

S → NP VP

NP → Det Nominal

NP → NP PP

A flight from Indianapolis to Houston on TWA

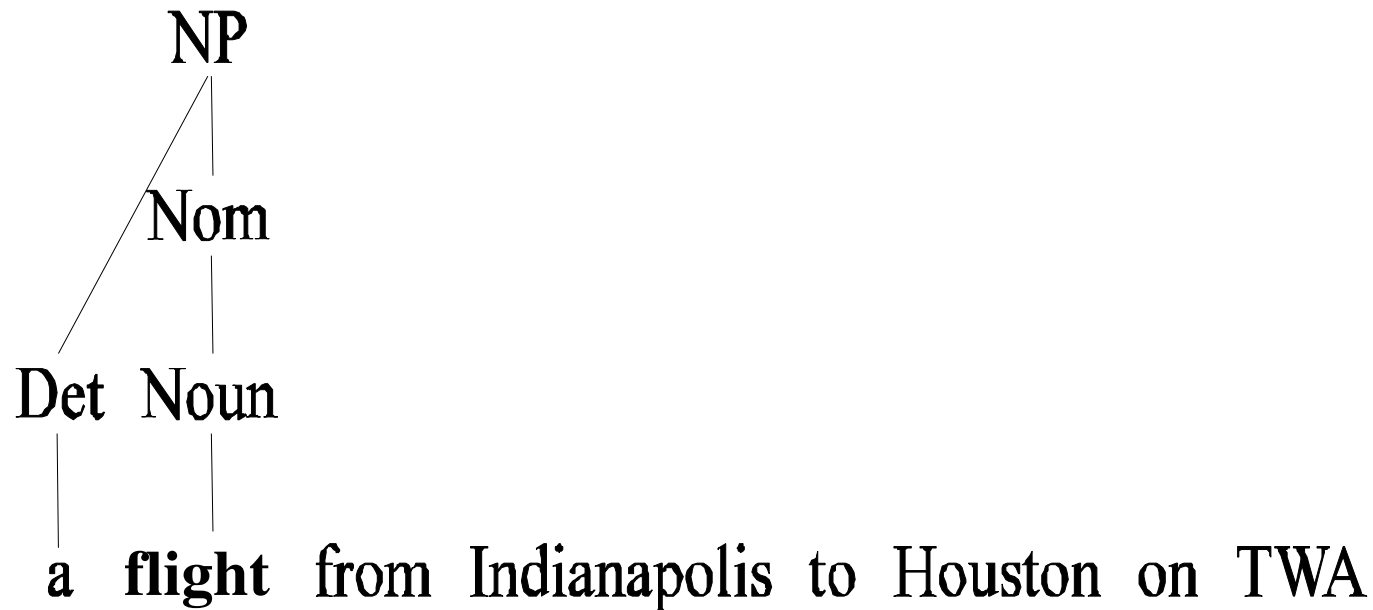
S -> Aux NP VP

S -> NP VP

NP -> Det Nominal

NP -> NP PP

A flight from Indianapolis to Houston on TWA



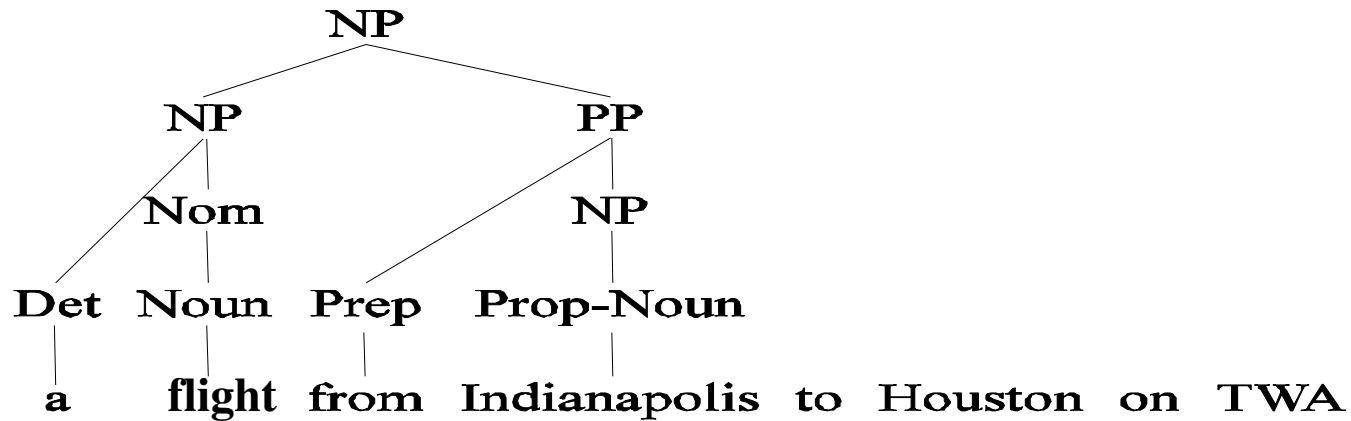
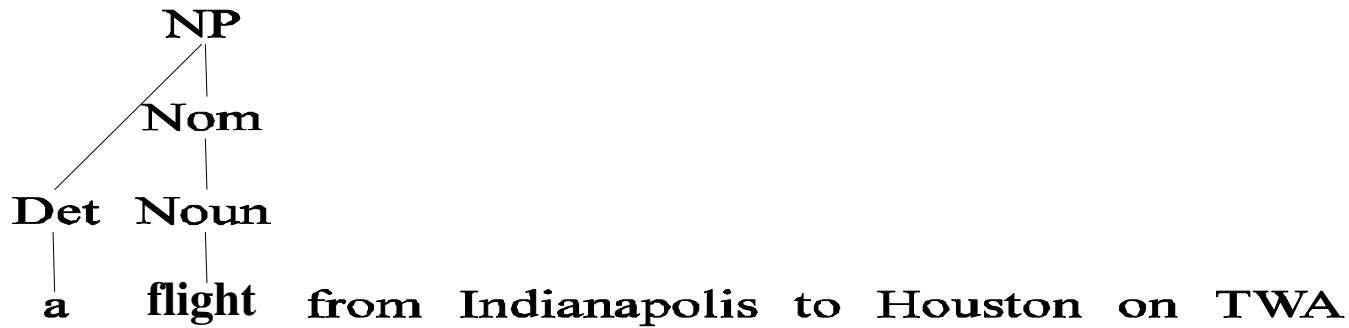
S -> Aux NP VP

S -> NP VP

NP -> Det Nominal

NP -> NP PP

A flight from Indianapolis to Houston on TWA



NP -> NP PP

NP
Nom
Det Noun

```

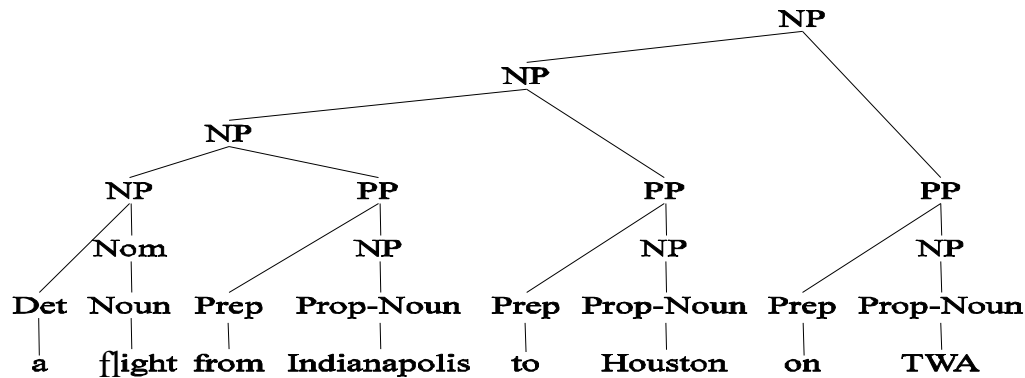
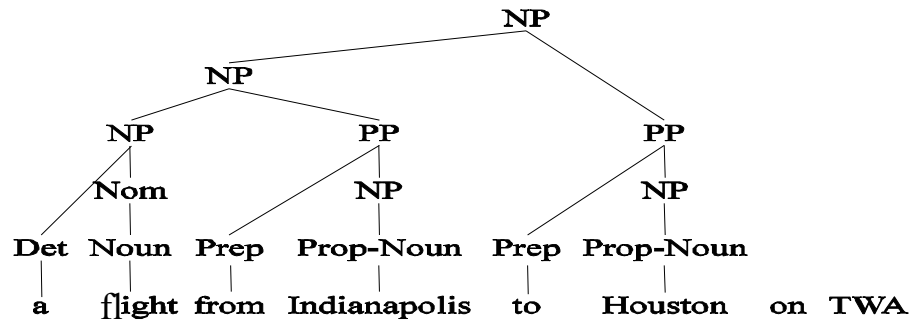
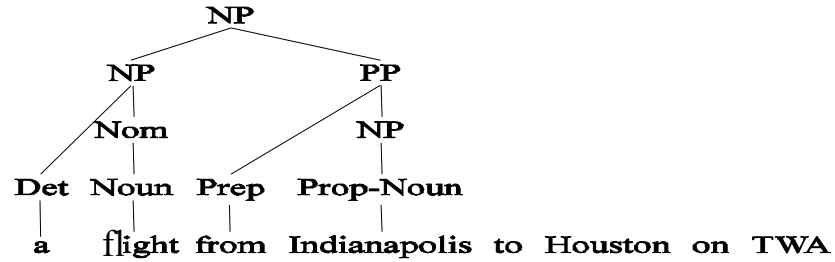
graph TD
    NP1[NP] --- NP2[NP]
    NP1 --- PP[PP]
    NP2 --- Nom[Nom]
    PP --- NP3[NP]
    Nom --- Det[Det]
    Nom --- Noun[Noun]
    NP3 --- Prep[Prep]
    NP3 --- PropNoun[Prop-Noun]
    Det --- a[a]
    Noun --- flight[flight]
    Prep --- from[from]
    PropNoun --- Indianapolis[Indianapolis]

```

Slide 43

NP -> NP PP

NP
Nom
Det Noun
a flight from Indianapolis to Houston on TWA



Aşağıdan-Yukarı Ayırıştırma (Bottom-Up Parsing)

$S \rightarrow NP VP$

$S \rightarrow Aux NP VP$

$S \rightarrow VP$

$NP \rightarrow Det Nominal$

$Nominal \rightarrow Noun$

$Nominal \rightarrow Noun Nominal$

$NP \rightarrow Proper-Noun$

$VP \rightarrow Verb$

$VP \rightarrow Verb NP$

$Det \rightarrow that \mid this \mid a$

$Noun \rightarrow book \mid flight \mid meal \mid money$

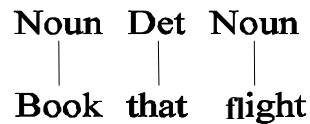
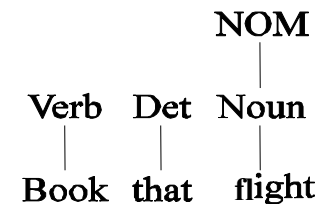
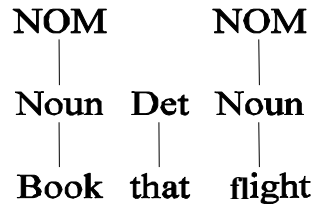
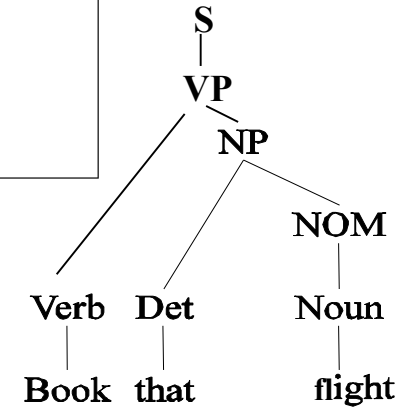
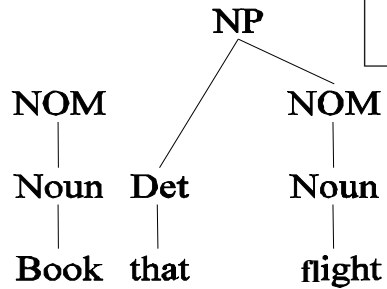
$Verb \rightarrow book \mid include \mid prefer$

$Aux \rightarrow does$

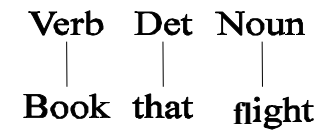
$Prep \rightarrow from \mid to \mid on$

$Proper-Noun \rightarrow Houston \mid TWA$

$Nominal \rightarrow Nominal PP$

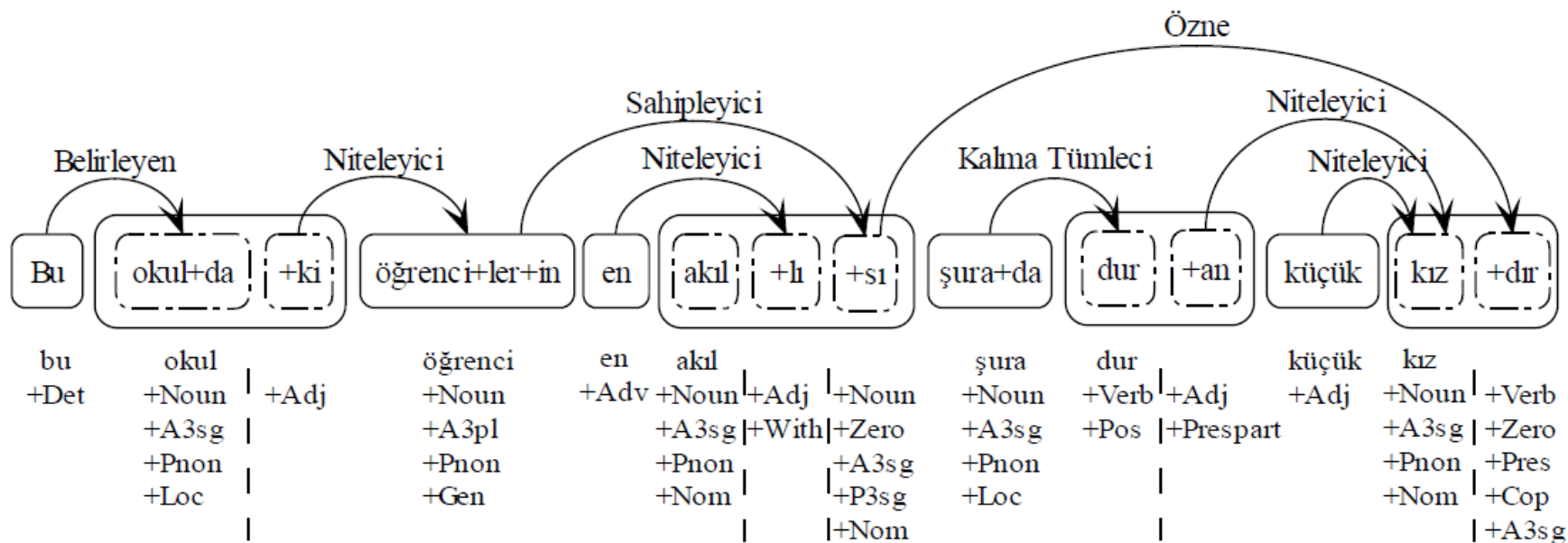


Book that flight



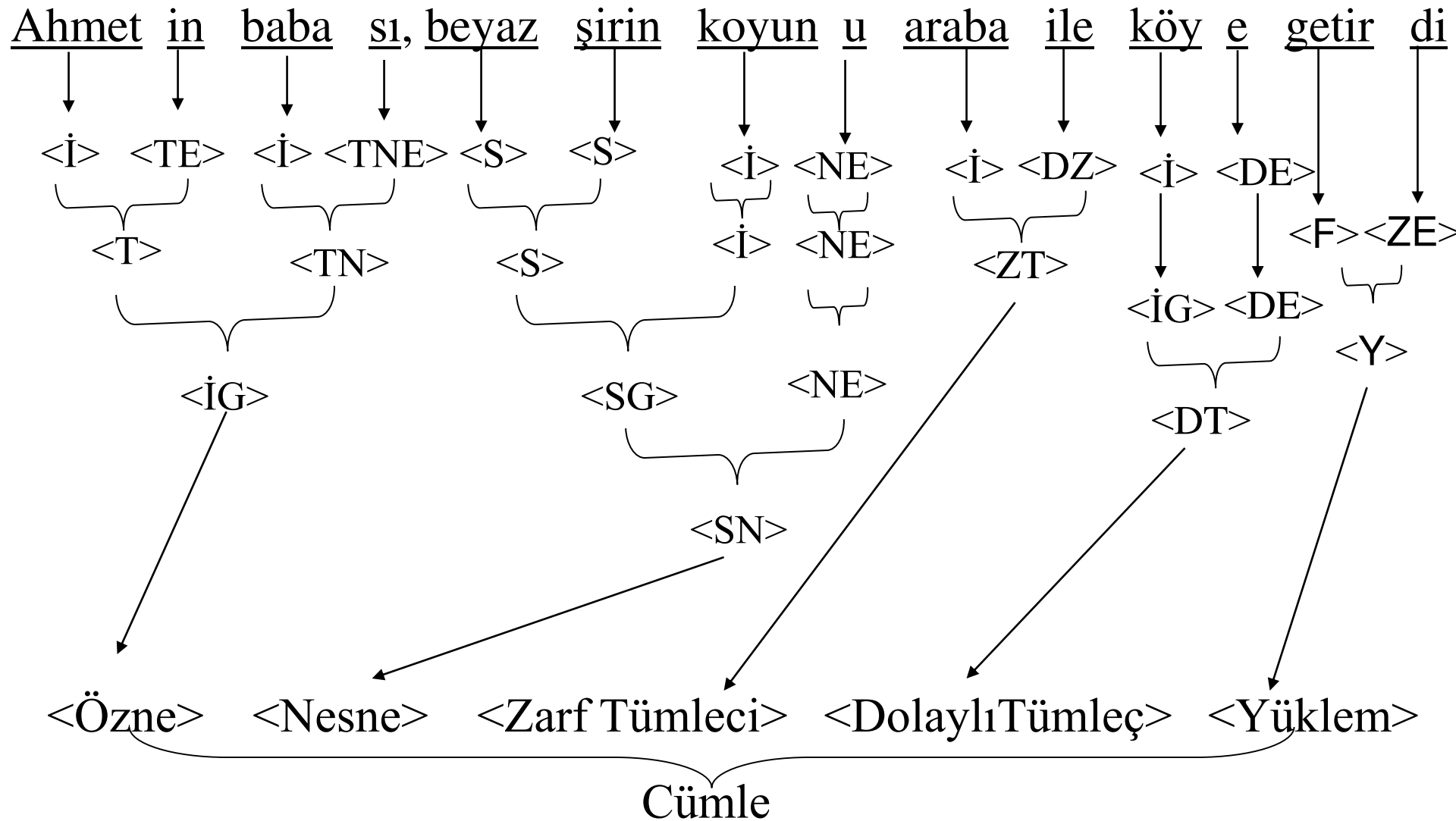
Bağılılık Ayırıştırması

- Bağımlılık çözümlemesi (dependency parsing) doğal dil işleme sürecinin sentaktik analiz aşamasında gerçekleştirilir.
- Cümlelerin yapısını analiz etmeye odaklanarak kelimeler arasındaki dilbilgisel ilişkileri belirlemeyi amaçlar.
- Bağımlılık çözümlemesi özellikle bir cümledeki kelimeler arasındaki sentaktik bağımlılıkları belirlemeyi hedefler ve bu ilişkileri kelimeler arasında yönlendirilmiş bağlantılar olarak temsil eder.
- Metnin sentaktik yapısını ve anlamını anlamada yardımcı olur, bilgi çıkarma, makine çevirisi ve duygu analizi gibi birçok NLP görevi için önemlidir.



Det: Belirleyen, Noun: İsim, Adj: Sıfat, Adv: Belirteç, Verb: Eylem, A3sg: 3. tekil kişi eki, A3pl: 3. çoğul kişi eki, Pnon: iyelik eki yok, Loc: -de hali, Gen: sahip olma hali, Nom: yalın hali, With: -li isimden sıfat türetme eki, Zero: ek almadan türetme, Pos: olumlu, Prespart: Şimdiki zaman ortacı, Pres: Şimdiki zaman, Cop: Koşaç

Cümlelerin Hiyerarşik Yapısı



"Ahmet'in babası beyaz şirin koyunu araba ile köye getirdi."

Cümlenin çözümlenmesi:

1. Cümle durumundan başlanır.
2. Cümleden bir sonraki duruma geçmenin şartı <özne> isimli duruma ait alt GGA'dan başarıyla dönülmesidir. Bu nedenle <özne>ye gidilir.
3. <özne> durumundan bir sonraki duruma geçmenin şartı kelimelerin <isim grubu> veya <sıfat grubu> oluşturmasıdır. "Ahmet": isim, "in": tamlayan_eki, "baba": isim, "sı": tamlanan_eki olduğu tanımlanmış olan veritabanından anlaşılır. Bu kelimelerden oluşan "Ahmet'in babası" kelime grubu <isim grubu> alt GGA'na uygundur. "Ahmet'in babası" kelime grubu <isim grubu> olarak atanır ve sonraki duruma geçilir.
4. Diğer durumlara geçişin bir çok farklı şartı vardır. Bu şartlardan herhangi biri sağlandığında diğer durumlara geçilecektir. "beyaz şirin koyunu" kelime grubunda; "beyaz": sıfat, "şirin": sıfat, "koyun": isim olduğu veritabanından anlaşılır ve bu üç kelime <sıfat grubu> testini başarıyla geçer. "koyunu" kelimesindeki "u" nesne_eki tanımına uygundur. "beyaz şirin koyun" kelime grubu, "u" ekiyle birleşince <sıfat nesnesi> tanımına uyar. <nesne> testini geçmenin şartlarından biri de <sıfat nesnesi> olarak tanımlandığından "beyaz şirin koyunu" kelime grubu <nesne> olarak atanır ve bir sonraki duruma geçilir. Bir sonraki durum yüklem, zarf tümleci veya dolaylı tümleç olmalıdır.

5. "araba" kelimesi <isim> olarak tanımlanan kelimeler arasındadır. <isim grubu> olmanın şartlarından biri de <isim> olmaktır. "ile" kelimesi <diğer_zarflar> arasında olduğundan "araba ile" kelime grubu <zarf tümleci> testini başarıyla geçer ve "araba ile" kelime grubu <zarf tümleci> olarak atanır. Bu durumdan sonra GGA'na göre yüklem veya dolaylı tümleş gelmelidir.

6. "köy" kelimesi <isim> olarak tanımlanan kelimeler arasındadır. "e" eki ise <dolaylı tümleş_eki> dir. Dolayısıyla "köye" kelimesi <dolaylı tümleş> olarak tanınır ve atanır. Bir sonraki durum GGA'na göre <yüklem> olmalıdır.

7. "getir" kelimesi veritabanına göre <fiil> dir. "di" eki de <zaman_eki> olarak tanımlandığından "getirdi" kelimesi <yüklem> olarak kabul edilir ve atanır.

8. Böylelikle cümlelerin unsurları olan <özne>, <nesne>, <zarf tümleci>, <dolaylı tümleş>, <yüklem> sırasıyla kabul edilmiş olur ve bu unsurlardan oluşan yapı cümle olarak kabul edilir ve atanır.

Genişletilmiş Geçiş Ağları – GGA

Augmented Network Transition Grammer- ATN

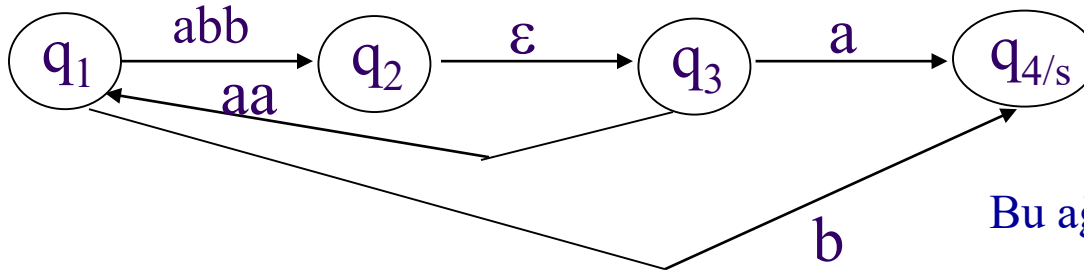
- Biçimsel dillerin, doğal dilleri tam olarak ifade edemediği durumlarda Genişletilmiş Geçiş Ağları tercih edilir.
- GGA, cümlelerin sözdizimsel analizinde kullanılır.
- GGA sonlu makinelere benzeyen durum ve bu durumlar arası geçiş kurallarından oluşmaktadır.

Üç bileşenden oluşur:

- En az bir başlangıç ve son durumu olan sonlu sayıdaki durumlar kümesi
- Belli bir metindeki mümkün olan harflerden oluşan küme
- Sonlu sayıdaki bir durumdan diğer bir duruma geçişi sağlayan geçişler kümesi

Genişletilmiş Geçiş Ağı örneği

- ❖ Geçiş ağlarında bir durumdan diğer bir duruma geçmek için gerekli harf okunur ve bu harf geçilecek olan duruma geçmek için gereken harfle karşılaştırılır. Uygun olması durumunda geçilir.
- ❖ Geçiş ağlarında doğru bir yol, bir başlangıç durumundan başlayıp, son duruma ulaşan geçişler sağlandığında tamamlanır.
- ❖ Geçişlerdeki harflerin birbirine eklenmesiyle oluşan metin, ağın kabul etmesi için verilen metinle aynı ise, bu metin ağ tarafından kabul edilmiş demektir.



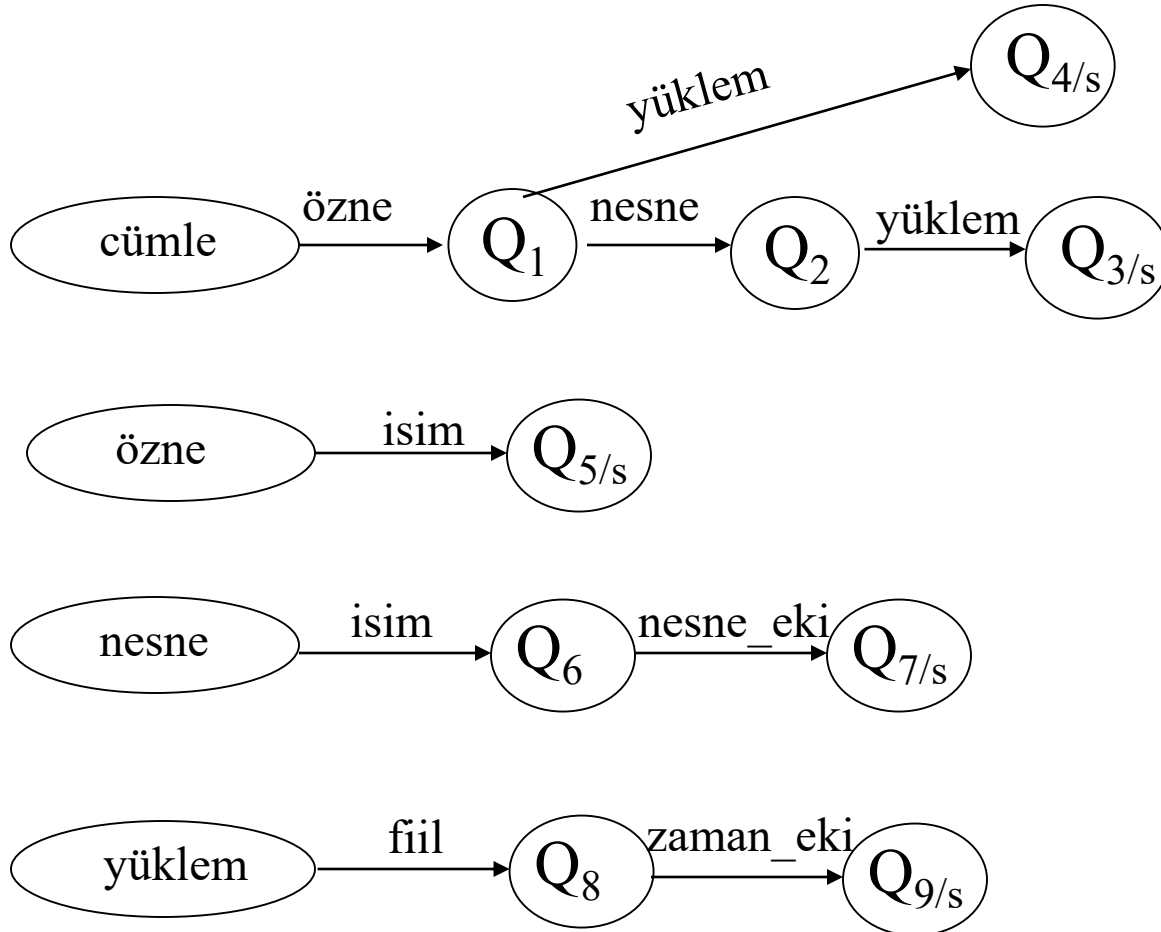
Bu ağda tanınabilecek birkaç metin örneği:

abba, abbbaabba, abbaab, b

Fakat *abbab, baab* yi tanıyamaz.

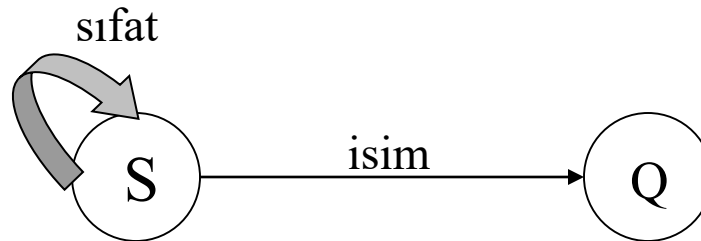
Basit bir Türkçe sözdizimi kural kümesinin GGA yapısı

“Ali camı kırdı” cümlesi morfolojik analiz sonucunda “Ali cam + ı kır + dı”



-
- ❖ Bir sözdizimi kural kümesinin GGA biçiminde ifade edilmesin yararı, cümle birimlerinin bir kez tanımlandıktan sonra sonsuz kez kullanılıyor olmasıdır.
 - ❖ Bir sıfat tamlaması bir kez tanımlanıp birçok yerde kullanılabilir.
 - ❖ GGA'da döngüler bulunabilir. Bu özellik ile dilin kabul ettiği cümle yapıları sayısı sonsuza kadar artırabilir.

“yırtık kırmızı büyük top”



Özetle...

- ❖ Bağlılık ayrıştırması, doğal dil işleme sürecinde "Sözdizimsel Analiz" aşamasında yer alır.
- ❖ Bu aşamada, cümlenin yapısı incelenir ve sözcükler arasındaki ilişkiler belirlenir.
- ❖ Bağlılık ayrıştırması, bir cümlenin içindeki her bir sözcüğün diğerleriyle olan ilişkisini belirler ve bu ilişkileri ağaç yapısı veya benzeri bir yapıda görselleştirir.
- ❖ Bu aşama, cümlenin anlamını daha iyi anlamak ve daha karmaşık dilbilgisel analizler için temel oluşturmak için gereklidir.