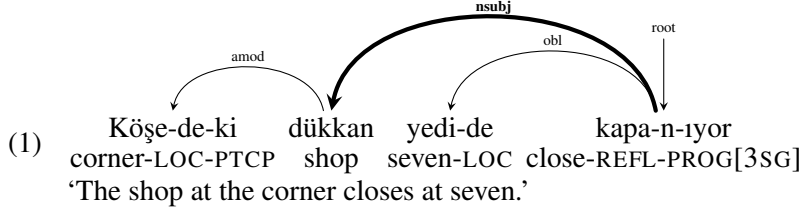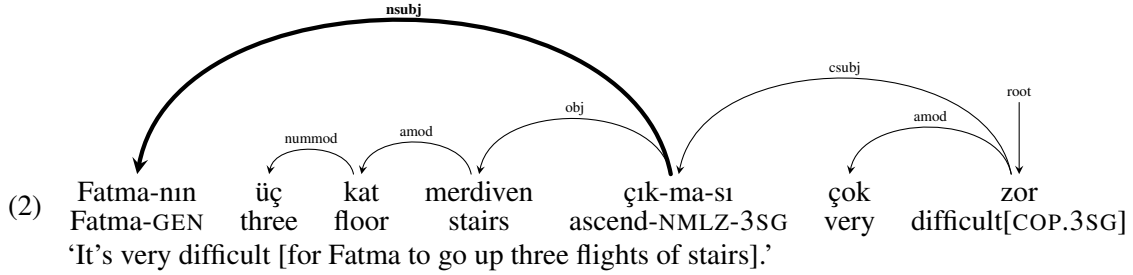# Suggested UD Guidelines for syntactic dependencies in Turkish

Utku Türk, Furkan Atmaca, Şaziye Betül Özateş, Gözde Berk, Seyit Talha Bedir, Abdullatif Köksal, Balkız Öztürk Başaran, Tunga Güngör, Arzucan Özgür
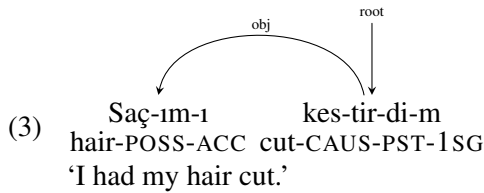
nsubj: nsubj marks the overt nominal subjects. Subjects in matrix clauses and some embedded clauses bear nominative case in Turkish represented by a null suffix (Göksel, Kerslake, 2004). In Example 1[1], the phrase *Köşedeki dükkan* is the subject of the sentence. *dükkan* is the head of this phrase. Therefore, it is tied to the predicate. As for *köşedeki*, it is a modifier (amod) for *dükkan*.

(1)

Köşe-de-ki	dükkan	yedi-de	kapa-n-ıyor
corner-LOC-PTCP	shop	seven-LOC	close-REFL-PROG[3SG]

'The shop at the corner closes at seven.'

Subjects in possessive embedded clauses bear genitive case marked by *-(n)In*. In UD, we do not use the nsubj relation to indicate subjects of Exceptional Case Marking (ECM) structures, those subjects are indicated by obj and depended on the upper clause predicate. In Example 2, the bracketed part is a nominalized embedded sentence. In these sentences, the predicate and the subject agree; we recognize the subject by genitive marker and the nominalized predicate by possessive marker.
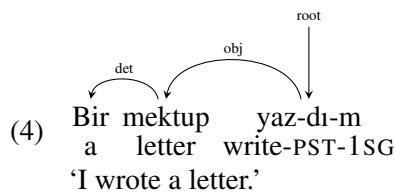

(2)

Fatma-nın	üç	kat	merdiven	çık-ma-sı	çok	zor
Fatma-GEN	three	floor	stairs	ascend-NMLZ-3SG	very	difficult[COP.3SG]

'It's very difficult [for Fatma to go up three flights of stairs].'

obj: obj marks overt nominal direct objects. Direct objects bear accusative case in Turkish. Accusative case is realized two ways depending on definiteness: *–(y)I* marks definite accusative arguments and null suffix marks indefinite accusative arguments. In Example 3, the phrase [*Saçımı*] is the direct internal argument of *kestir-* 'get sth cut'.

(3)

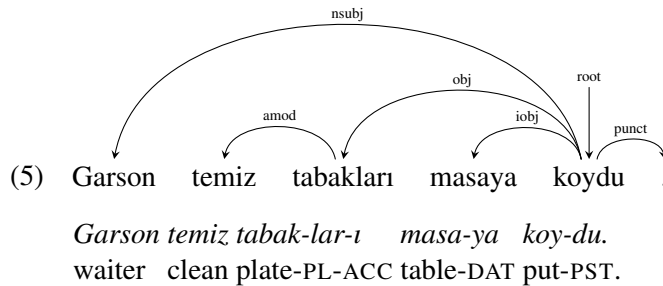| Saç-ım-ı	kes-tir-di-m
hair-POSS-ACC	cut-CAUS-PST-1SG

'I had my hair cut.'

In Example 4, the unmarked phrase [*mektup*] is the direct internal argument of *yaz-* 'to write'. It is a potential candidate for being a nominal subject as well; however, it is eliminated by the fact that the predicate is inflected by 1st person. If the predicate were to be inflected to 3rd person like *Bir mektup yazdı* 'He wrote a letter.' then we are supposed to look for semantic clues and imagine probable contexts. In this case semantic clues would be enough since a letter cannot write as the predicate's external argument demands agency.

---

[1]Abbreviations used in this paper are as follows: 1 = first person, 2 = second person, 3 = third person, ABL = ablative, ACC = accusative, ADJ = adjective, ADV = adverbial, CAUS = causative, COND = conditional, COP = copula, DAT = dative, FUT = future, GEN = genitive, INF = infinitive, INS = instrumental, LOC = locative, NEG = negative, NMLZ = nominalizer, PL = plural, POSS = possessive, PROG = progressive, PRS = present, PST = past, PTCP = participle, REFL = reflexive, SG = singular.
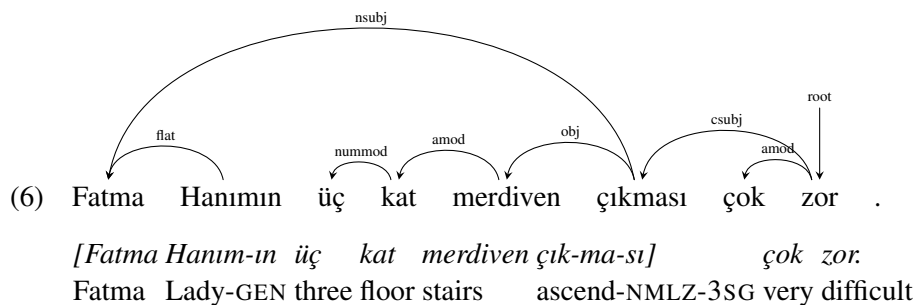
(4)

```
          root
        ┌──┘
    obj │
  ┌──┘  │
det│    │
```

Bir mektup yaz-dı-m
a   letter   write-PST-1SG

'I wrote a letter.'

iobj: iobj marks overt nominal indirect objects of ditransitive predicates. In Turkish they typically marked by dative case which is marked by the suffix –(y)A. In UD iobj is only allowed when there is an overt direct object.

(5)
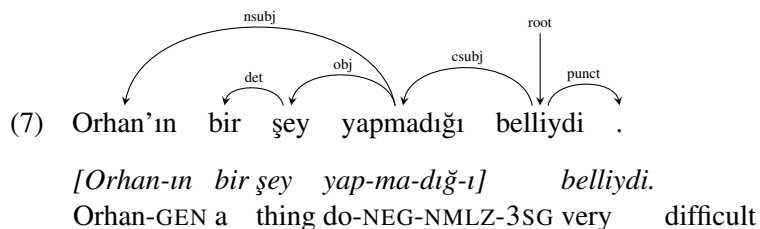
Garson temiz tabakları masaya koydu .

Garson temiz tabak-lar-ı   masa-ya koy-du.
waiter   clean plate-PL-ACC table-DAT put-PST.

The waiter put the clean plates down on the table.

In this example the NP [temiz tabakları] is the direct object and is obj>koydu, whereas the NP [masaya] is the indirect NP and iobj>koydu.
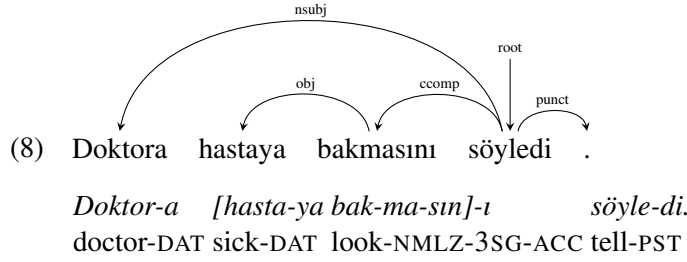
csubj: Clausal subject arguments are generally marked *-mA, -mAK, -DIK, -(y)AcAK* nominalizers. They are nominative in matrix clauses and some embedded clauses, genitive in possessive embedded clauses.

(6)

Fatma Hanımın üç kat merdiven çıkması çok zor .

[Fatma Hanım-ın üç   kat merdiven çık-ma-sı]   çok zor.
Fatma Lady-GEN three floor stairs   ascend-NMLZ-3SG very difficult

It's very difficult [for Fatma Hanım to go up three flights of stairs].

In this example, the CP [Fatma Hanım-ın üç kat merdiven çıkması] has the head [çıkması] as VP, which is nominalized via *-mA* and genitive-possessive construction.
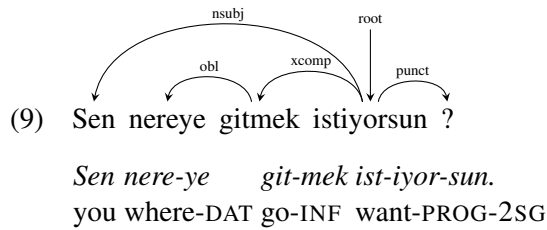
(7)

Orhan'ın bir şey yapmadığı belliydi .

[Orhan-ın bir şey yap-ma-dığ-ı]   belliydi.
Orhan-GEN a   thing do-NEG-NMLZ-3SG very   difficult

It was obvious [that Orhan wasn't doing/hadn't done anything].

ccomp: Clausal complements are any infinite or finite clauses that are direct objects depended on a predicate. When they are infinite, they are expected to bear an overt case, if not we prefer to use xcomp.

(8)

Doktora   hastaya   bakmasını   söyledi   .

Doktor-a   [hasta-ya bak-ma-sın]-ı   söyle-di.
doctor-DAT sick-DAT look-NMLZ-3SG-ACC tell-PST

S/he told the doctor [to check the patient].

In this example, the VP [**bakmasını**] is the head of the CP [*hastaya bakmasını*] and is `ccomp`>söyledi because it is the clausal complement of the predicate. Notice that this CP has a covert *PRO$_i$* that is coindexed with another element that occurred before: *Doktora$_i$*. We understand existence of this invisible subject of the embedded CP via the 3rd person agreement marker on the predicate [*bakma-**sın**-ı*].
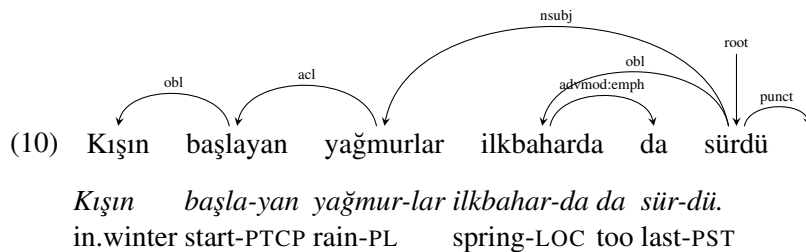
xcomp: xcomp is used for the clausal infinite complements that lack subject place (subject of the embedded clauses is controlled by the subject of the matrix clause) and they are not marked with any case marker. Typically, these are embedded predicates that suffixed by *-mAK* under the matrix predicate *iste-*.

PS: In UD, xcomp is used for all open clauses without subject. We however, put additional condition to use this relation: the embedded clause must be caseless.

(9)

Sen   nereye   gitmek   istiyorsun   ?

Sen nere-ye   git-mek ist-iyor-sun.
you where-DAT go-INF want-PROG-2SG
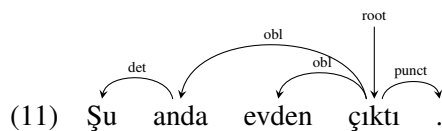
Where do you want to go?.

In this example, the embedded CP [*nereye gitmek*] lacks any subject and case marking. It is also the complement of the matrix predicate *istiyorsun*. Therefore, the head of the embedded CP *gitmek* should be `xcomp`>istiyorsun.

obl: obl is used to mark non-core elements in a sentence that can indicate time, place, manner etc. They can be marked by a grammatical case (many time obliques are not marked with any case), or dependency `case`.

(10)

Kışın   başlayan   yağmurlar   ilkbaharda   da   sürdü   .

Kışın   başla-yan yağmur-lar ilkbahar-da da sür-dü.
in.winter start-PTCP rain-PL   spring-LOC too last-PST

The rain that began in the winter continued in the spring, too.

In this example, oblique objects *kışın* and *ilkbaharda* indicate the time, former of which is a dependent to the embedded predicate, the latter to matrix. Generally, any nominal that indicates time of the verb will be marked with `obl`.

(11)

Şu   anda   evden   çıktı   .

Şu an-da   ev-den   çık-tı.
that moment-LOC house-ABL get.out-PST

S/he has just left the house.

In this example, *evden* is a typical example of non-core argument (adjunct). The predicate *çık-* is an unergative verb that cannot have an internal argument. Besides accusative, all cases could potentially be a sign of an adjunct in Turkish.
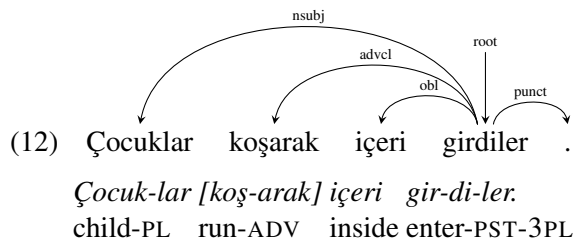
vocative: Indicates vocative elements in a sentence. They directly depend on the predicate of that clause.

expl: N/A

dislocated: `dislocated` is utilized in Turkish when a sentence contains an element that are exemplified usually after a colon.
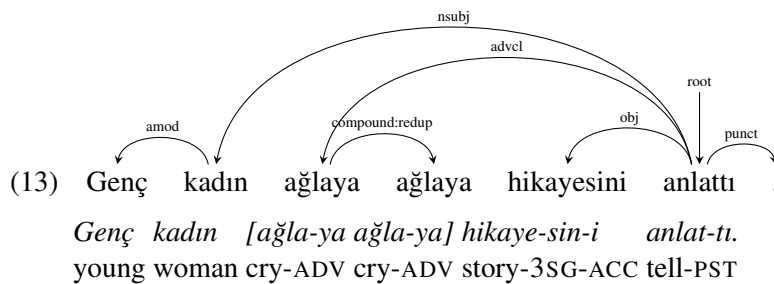
advcl: `advcl` is the only non-core clausal element. It basically marks the predicate of an embedded clause. They are identified by some items like *olarak* or *için* marked by case or such suffixes: *-(y)ArAK, -DIkçA, -mAdAn, -DIğIndAn, -ken, -IncA, mAktAnsA* etc.

`advcl:cond` is used to tie heads of conditional adverbial clauses to the upper clause.

(12) Çocuklar   koşarak   içeri   girdiler   .

Çocuk-lar [koş-arak] içeri   gir-di-ler.
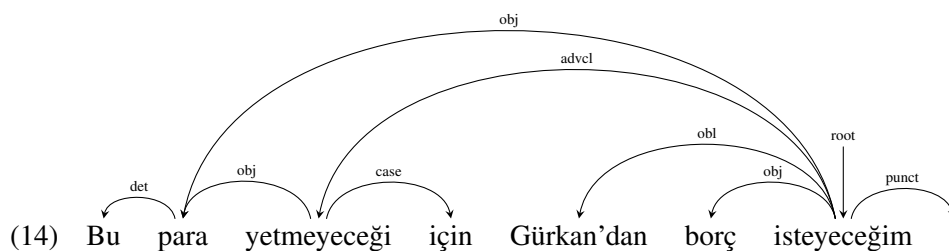child-PL   run-ADV   inside enter-PST-3PL

The children came in [running].

In this example, the adverbial phrase in brackets is a deverbal adverb. Note that many adverbial clauses naturally lack subject. (11) and (12) are examples of this sort. These adverbial clauses have a PRO$_i$ that is coindexed with the subject of the relevant clause. However, some adverbial clauses like (13) may have a subject, when examples like (13) lack overt subject their subject is assumed to be identical with the subject of the upper clause.

(13) Genç   kadın   ağlaya   ağlaya   hikayesini   anlattı   .

Genç kadın   [ağla-ya ağla-ya] hikaye-sin-i   anlat-tı.
young woman cry-ADV cry-ADV story-3SG-ACC tell-PST
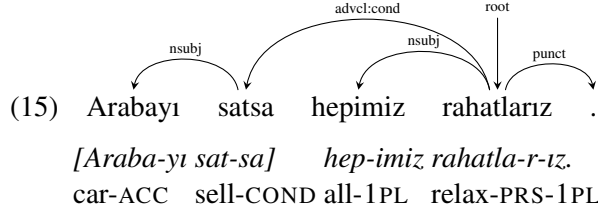
The young woman told her story [weeping (continually)].

In this example we see a compound adverbial clause. Duplication is a fairly common operation in Turkish for both denominal and deverbal adverbials.

(14) Bu   para   yetmeyeceği   için   Gürkan'dan   borç   isteyeceğim   .

[Bu para   yet-me-yeceğ-i   için]   Gürkan-dan borç   iste-yeceğ-im
this money suffice-NEG-PTCP-3SGfor Gürkan-ABL debt   want-FUT-1SG

[As this money won't be enough] I'm going to ask Gürkan a loan.

In this example the embedded sentence is an adverbial clause with a subject NP [Bu para], so its predicate *yetmeyeceği* is inflected to 3rd person which is different than the matrix subject which is an overt 1st person.
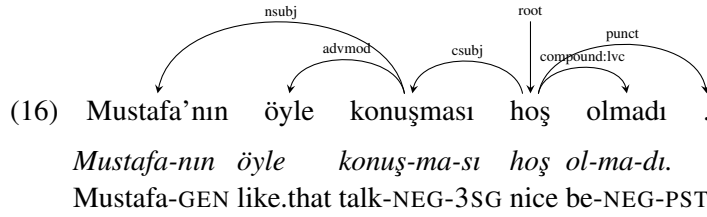
(15)

Arabayı   satsa   hepimiz   rahatlarız   .

[Araba-yı sat-sa]   hep-imiz rahatla-r-ız.
car-ACC   sell-COND all-1PL   relax-PRS-1PL

[As this money won't be enough] I'm going to ask Gürkan a loan.

In this example, *satsa* is the head of the conditional phrase. These are regarded as a sort of adverbial clause in UD. In Turkish, they are marked by *-sA* suffix.
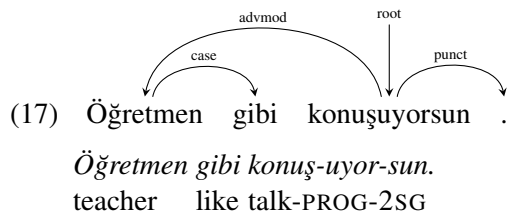
advmod: advmod is used to mark adverbs that modify an adjective, a verb or another adverb in terms of manner, degree, or frequency. It is also used to mark sentence level adverbials like *özellikle, ayrıca, bilhassa, buna ilaveten* etc.

advmod:emph is used to mark the clitic *dA* which indicates inclusion (means like 'also, too') and appears in some fixed variations like *ya da, hem de, ne de* etc. *bile* is also marked with this relation.
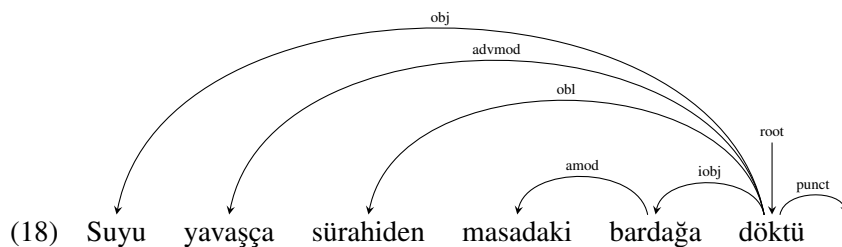
(16)

Mustafa'nın   öyle   konuşması   hoş   olmadı   .

Mustafa-nın öyle   konuş-ma-sı hoş ol-ma-dı.
Mustafa-GEN like.that talk-NEG-3SG nice be-NEG-PST

It was unpleasant that Mustafa talked like that.

In this example *öyle* 'like that' is the adverbial modifier of the embedded predicate.

(17)

Öğretmen   gibi   konuşuyorsun   .

Öğretmen gibi konuş-uyor-sun.
teacher   like talk-PROG-2SG

You're talking like a teacher.

The postpositive particle *gibi* 'like, as' is understood to be head of the postpositional phrase [PP Öğretmen gibi]; in UD we understand it as merely a dependent (marked by case) of a nominal which it makes into a adverbial modifier.

(18)

Suyu   yavaşça   sürahiden   masadaki   bardağa   döktü   .

Su-yu   yavaş-ça sürahi-den masa-da-ki   bardağ-a dök-tü.
water-ACC slow-ADV jug-ABL   table-LOC-ADJ glass-DAT pour-PST

S/he slowly poured the water from the jug into the glass.

-cA is a typical adverbializer in Turkish, which is mostly added to the adjectives.

(19) Oya işe gitti Ali de .

Oya iş-e   git-ti, Ali de
Oya work-DAT go-PST Ali too

Oya went to work, and Ali, too.

In this example, emphatic *dA* gives meaning of inclusion to Ali and connects it to the previous predicate.

discourse: discourse is used to mark discursive elements such as exclamations, murmurs, and other semantically vacuous sounds. *Evet* 'yes' and *hayır* 'no' are also marked with discourse. These elements depend on the predicate of the relevant clause.

(20) Ambülans çabuk geldi çok şükür .

Ambülans çabuk gel-di   [çok   şükür].
ambulance quick come-PST thankfully

Thank god, ambulance came quickly.

In this example, *çok şükür* is a compound that expresses thankful wishes. In (21)[2] (Göksel, Kerslake, 2004) provides some list of these discursive elements like this:

(21)   *inşallah* 'God willing, hopefully' *umarım* 'I hope' *Allahtan/bereket versin* 'fortunately' *çok şükür* 'thank goodness, fortunately' *iyi ki* 'it's good idea that...' 'thank goodness' *maalesef/ne yazık ki* 'unfortunately' *tabii (ki)/doğal olarak* 'of course, naturally'

aux: Turkish lacks most auxiliary elements, the only element we utilize in UD is aux:q for the question clitic *mI*.

(22) Nermin okula mı gitmiş ?

Nermin okul-a   mı git-miş?
Nermin school-DAT Q go-PST

Has Nermin gone to school?

(23) Nermin   okula   gitmiş   mi   ?

---

[2]Abbreviations used in this paper are as follows: 1 = first person, 2 = second person, 3 = third person, ABL = ablative, ACC = accusative, ADJ = adjective, ADV = adverbial, CAUS = causative, COND = conditional, COP = copula, DAT = dative, FUT = future, GEN = genitive, INF = infinitive, INS = instrumental, LOC = locative, NEG = negative, NMLZ = nominalizer, PL = plural, POSS = possessive, PROG = progressive, PRS = present, PST = past, PTCP = participle, REFL = reflexive, SG = singular.
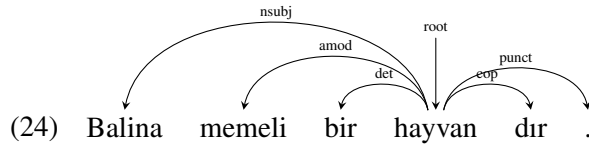
Nermin okul-a   git-miş mi?
Nermin school-DAT go-PST Q

Has Nermin gone to school?

In example (22), question particle *mI* focuses over the adjunct [NPokula] therefore it is dependent on it whereas in (23) it focuses on the predicate.

cop: `cop` is used to mark copular elements that are either attached or cliticized to the predicate. *-DIr, -(y)DI, -(y)mIş, -(y)sA, -(y)ken* are considered as Turkish copulas and are always marked `cop`.
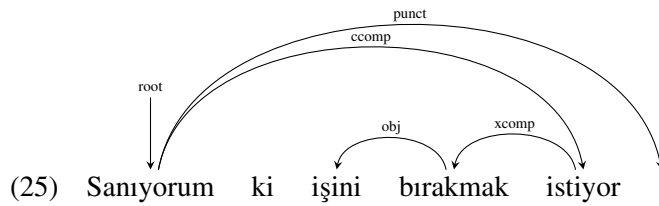
(24) Balina   memeli   bir   hayvan   dır   .

Balina memeli   bir hayvan-dır.
Whale mammal a   animal-COP

Balina is a mammal.

-DIr is the only copular element in Turkish that solely comes in suffixal form, others may turn into clitics as well: *-(y)DI ∼ idi, -(y)mIş ∼ imiş, -(y)sA ∼ ise, -(y)ken ∼ iken.*

mark: `mark` is generally used to mark *ki* that conjoins two sentences.

(25) Sanıyorum   ki   işini   bırakmak   istiyor   .

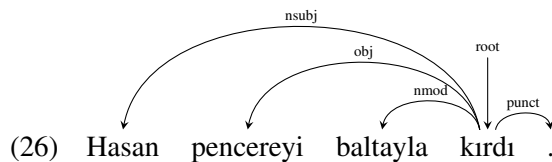*San-ıyor-um   ki iş-in-i   bırak-mak   ist-iyor.*
suppose-PROG-1SG that job-3SG-ACC leave-NMLZ want-PROG

I think (that) s/he wants to leave his/her job.

nmod: Bare `nmod` is the relation that indicates instrumental nouns.

`nmod:poss` is an extremely common relation that marks a possessive dependent that modifies possessed noun. Possessive dependent might bear -In genitive suffix or not. Note that for the dependents nominalized embedded clauses whose structure looks a lot like a possessive noun phrase we use `nsubj` rather than `nmod:poss`.

PS: In UD, `nmod` is supposed to modify a noun head. We, however, use `nmod` in a particular situation that modifies a predicate.

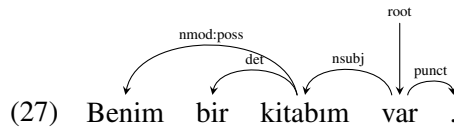(26) Hasan   pencereyi   baltayla   kırdı   .

Hasan pencere-yi   balta-yla   kır-dı.
Hasan window-ACCaxe-INS break-PST

Hasan broke the window with (an) axe.

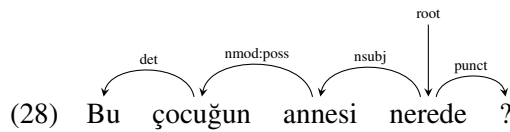In this example, the noun phrase *baltayla* is in instrumental case and modifies a predicate.

(27)

```
        nmod:poss
              det    nsubj   root
                              punct
    Benim   bir   kitabım   var    .
```

Ben-im bir kitab-ım   var.
I-GEN   a   book-1SG there.is

I have a book.

Turkish lacks the verb *have to* indicate possession. So, it uses possessive existential sentences like (26) to indicate possession. That is, in order to say *x has y*, it says *x's y exists*. In these structures, we ought to use `nmod:poss` to indicate that y belongs to x.
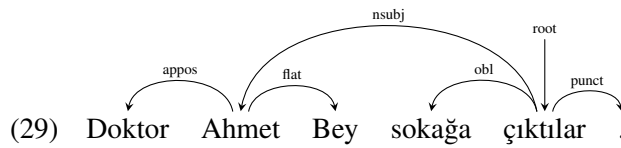
(28)

```
     det    nmod:poss   nsubj   root
                                 punct
    Bu   çocuğun   annesi   nerede   ?
```

Bu çocuğ-un anne-si   nerede?
this child-GEN mother-3SG where

Where is this child's mother?

appos: `appos` is a relation between two nominals, on being the explanation of the other. We mark the modifying nominal by this relation. Modifying nouns generally indicate the title, relation. If the explanation is in parentheses `appos` is still applicable.
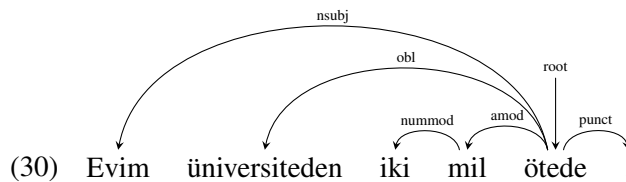
(29)

```
                nsubj
      appos           obl    root
             flat              punct
    Doktor  Ahmet  Bey  sokağa  çıktılar  .
```

Doktor Ahmet Bey sokağ-a   çık-tı-lar.
doctor Ahmet Mr street-LOC get.out-PST-3PL

Mr Doctor Ahmet has gone out on the street.

nummod: Numerical modifiers are ordinal, cardinal numbers and numbers inflected with *-lArcA* like *yüzlerce, onlarca* marked by nummod. Numerical modifiers with multiple items are tied together via flat.
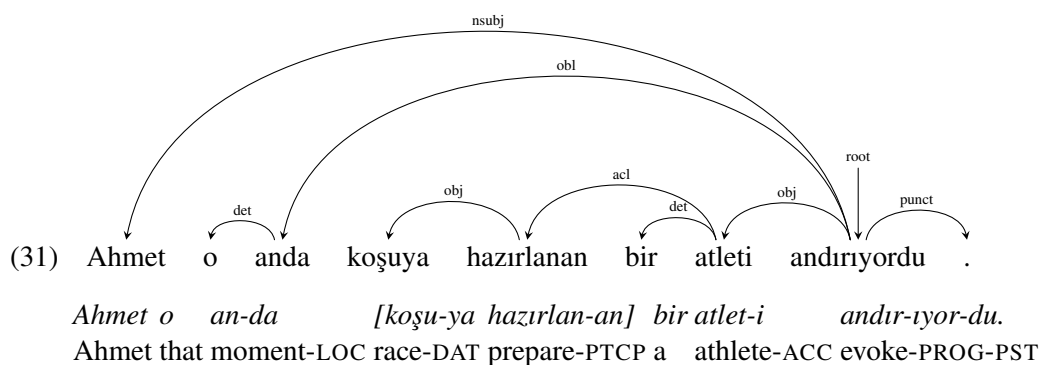
(30)

```
                nsubj
                     obl
                  nummod  amod  root
                                 punct
    Evim  üniversiteden  iki  mil  ötede  .
```

Ev-im   üniversite-den iki mil ötededir.
house-1SG university-ABL two mile is.away

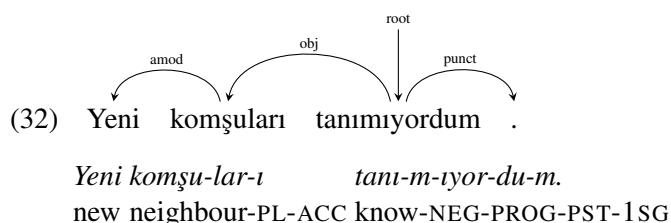My house is two miles away from the university.

acl: `acl` is a relation for adjectival clauses which are embedded structures that modify a noun. They are typically suffixed by *–(y)An, -DIK, -(y)AcAK, -Ar, -mIş, -AsI*. Another rather rare construction is via *ki* particle in a sentence like *Ali ki arkadaşımdır beni çok sever arkadaşımdır* is depended on *Ali* via `acl`.

(31) Ahmet o anda koşuya hazırlanan bir atleti andırıyordu .

Ahmet o an-da [koşu-ya hazırlan-an] bir atlet-i andır-ıyor-du.
Ahmet that moment-LOC race-DAT prepare-PTCP a athlete-ACC evoke-PROG-PST

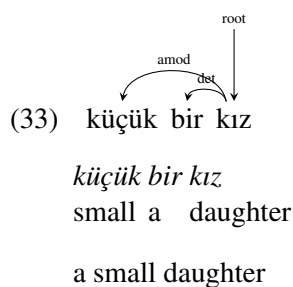At that moment Ahmet looked like an athlete [preparing for a race].

In this example, the CP [koşuya hazırlanan] modifies the DP [bir atleti]. *hazırlanan* is the predicate of the CP and bears one of the adjectivizer suffixes.

amod: amod is a relation that marks adjectival modifiers that modify the head of a nominal phrase. In Turkish many nouns are convertible in adjectives via conversion or suffixes like *-lI, -sIz, -cA*.

(32) Yeni komşuları tanımıyordum .

Yeni komşu-lar-ı tanı-m-ıyor-du-m.
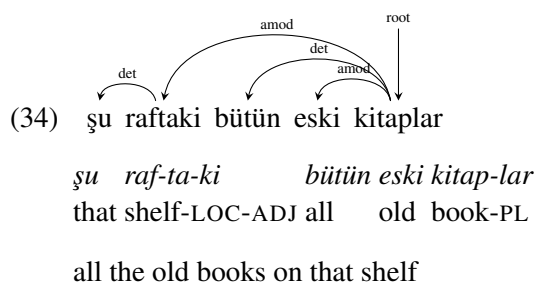new neighbour-PL-ACC know-NEG-PROG-PST-1SG
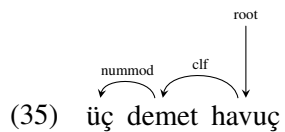
I didn't know the new neighbours.

det: det is a relation between a noun head and its determiner. Determiners in Turkish are demonstratives (*bu* 'this', *şu* 'that', *o* 'that'), indefinite article (*bir* 'a'), quantifiers (*bazı* 'some', *birkaç* 'a few', *hiçbir* 'none', *bütün* 'all', *tüm* 'every' etc.), or interrogative determiner (*hangi* 'which'). Possessive determiners (*benim* 'my', *senin* 'your' etc.) are, on the other hand, are marked by nmod:poss.

(33) küçük bir kız

küçük bir kız
small a daughter

a small daughter

In this example, we see a typical use of an indefinite determiner who is often found between an adjective and the head noun.

(34) şu raftaki bütün eski kitaplar

şu raf-ta-ki bütün eski kitap-lar
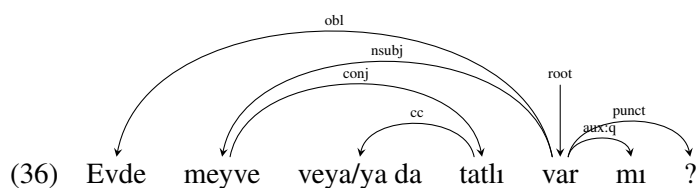that shelf-LOC-ADJ all old book-PL

all the old books on that shelf

clf: `clf` is a relation that marks classifiers in Turkish. They are rather rare and generally occur between a numeral and a noun head. *demet, baş, çay kaşığı, tutam, tane, fincan, kap, teneke* are some examples of classifiers in Turkish.

(35)   üç demet havuç

üç   demet havuç
three bunch carrot

three bunches of carrot

conj: The relation `conj` asymmetrically conjoins two syntactically identical elements, head being the first conjunct. Items could be conjoined via commas or coordinators such as *ve, veya, ya da* etc.

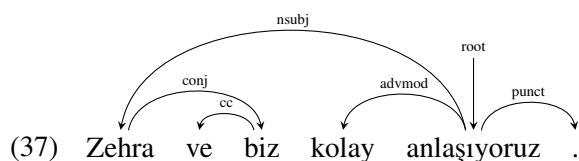(36)   Evde   meyve   veya/ya da   tatlı   var   mı   ?

Ev-de   meyve veya/ya da   tatlı   var mı?
house-LOC fruit   or   dessert there.is Q

Is there any fruit or any [sort of] sweet in the house?

As per UD rules, we connect second component to the first if there is no meaningful hierarchy between components.

cc: `cc` is utilized for the coordinators that coordinate two elements like *ve, veya* etc. However, some conjunctions (such as *ancak, çünkü*) may not contain two elements, they modify entire sentence. In that case their dependency is still `cc`, but they depend on the predicate.
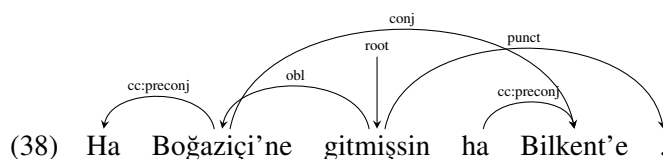
`cc:preconj` is a dependency that marks correlative conjunctions such as: *hem... hem (de), ne... ne (de), ya... ya (da), ha ha...*

(37)   Zehra   ve   biz   kolay   anlaşıyoruz   .

Zehra ve   biz kolay anla-ş-ıyor-uz.
Zehra and we easy   understand-REFL-PROG-1PL

We and Zehra get along easily(=well).

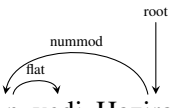As we see in this example, `cc` is connected to the next coordinated item in a conjunction.

(38)   Ha   Boğaziçi'ne   gitmişsin   ha   Bilkent'e   .

Ha   Boğaziçi-ne   git-miş-sin   ha   Bilkent-e.
be.it Boğaziçi-LOC go-PST-2SG be.it Bilkent-LOC

It doesn't matter whether you go to Boğaziçi or to Bilkent.

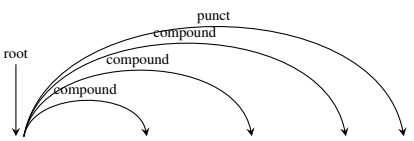Just like `cc`, `cc:preconj` is also linked to the next item.

fixed: **N/A**

flat: `flat` is the relation that ties together proper nouns and numerals when they have more than one element. We also use it to mark the items connected by a dash or a slash. In any case, first element is always the head.

(39)   on yedi Haziranda

on yedi Haziran-da
ten seven June-LOC

on the 17th of June'

compound: `compound` is the relation we use when a noun modifies a noun head to give information about its material. In this case, the head is the modified noun. It is also used when there is an idiomatic expression in the sentence, in which case the first element would be the head.
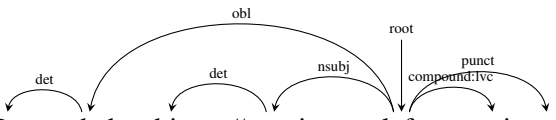
`compound:lvc` is the relation that marks light verbs. Light verbs in Turkish follow their main verb, making a compound. *et-, yap-, ol-, kıl-* are common light verbs in Turkish.

`compound:redup` is a relation used for reduplications like *hemen hemen, yavaş yavaş, ufak tefek, zar zor* etc. The first element is the head as usual.
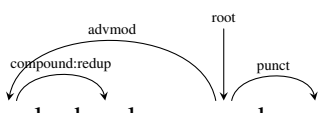
(40)   Hasan yağmurdan kaçarken doluya tutuldu .

Hasan   [yağmurdan kaçarken doluya tutuldu].
Idiomatic reading: 'Hasan   jumped out   of the frying pan into the fire.'

Literal reading: 'While Hasan was trying to escape the rain, he was seized by hail.'

In this example, we see a long idiomatic expression. No matter how long it is, all idioms are to be linked via `compound` and headed by the first component.

(41)   Bu   sabah   bir   öğrencim   telefon   etti   .

Bu sabah   bir öğrenci-m   telefon   et-ti.
this morning a   student-1SG telephone make-PST

A student of mine rang this morning.

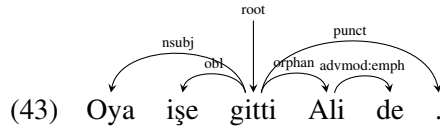(42)   Horul   horul   uyuyordu   .

Horul   horul   uyu-yor-du.
snoringly snoringly sleep-PROG-PST

11

'She was sleeping, snoring away.'

list: The relation `list` is used to indicate strings of phone numbers, post codes or full dates.

parataxis: We use `parataxis` in order to mark definitions, explanations and full sentences in parentheses.

orphan: `orphan` is utilized when a predicate is elided in a construction and one or more elements left with no overt head. In such cases, the element that is coordinated is promoted to the head of the clause; other elements are tied to this head via `orphan`.

(43)

Oya işe gitti Ali de .

Oya iş-e   git-ti,   Ali de
Oya work-DAT go-PST Ali too

Oya went to work, and Ali, too.

The sentence fragment Ali de is actually undergone ellipsis and the full form would be [$_{CP}$Ali de (gitti)]. *Ali*, being the subject, now lacks a predicate to be tied to. As a remedy, in UD, orphan is utilized to tie this particle onto an upper head that is relevant in the context. In this case, it is the predicate before since *Ali* is a subject of that.

goeswith: We use `goeswith` is utilized to tie together wrongly separated text.

reparandum: We use `reparandum` is utilized to mark speech repair.

punct: We use `punct` is utilized to mark speech repair.

root: `root` a special relation marks the roots of the sentence. Unlike other relations head of this relation is a hypothetical marker outside of the sentence. In UD, every sentence must contain one and only one `root` relation.

dep: `dep` is a last resort dependency when no other dependency is suited to use. Sometimes we use this dependency in order to indicate the numbers or letters as bullet heads like: *A) [SENTENCE]*, *a1-[SENTENCE]*.

References

Göksel Aslı, Kerslake Celia. Turkish: A comprehensive grammar. 2004.