

## Chapter 2

# Matters of size and deficient functional categories in three Turkic languages: Turkish, Turkmen, and Noghay

Meltem Kelepir


Boğaziçi University

This chapter compares and contrasts the verbal domain of the nominalized indicative complement clauses in three Turkic languages: Turkish, Turkmen, and Noghay, and argues for two points: (i) the size of this verbal domain under the nominalizing head is the same as the main clause verbal domain in Turkmen and Noghay, but smaller in Turkish, and (ii) regardless of the size of the verbal domain, in all these languages the highest inflectional category lacks certain morphosyntactic features (such as aspect, mood/modality, and tense, depending on the language) and the morpheme that heads this category is an underspecified morpheme, despite appearances.

## 1 Introduction

Turkic languages are known to be typical examples of languages that predominantly employ nominalization in subordinate clauses, and are, for instance, classified by Givón (2009) under “extreme nominalization languages” where all non-main clauses are nominalized to some degree. Givón reports that the following properties are the three most conspicuous telltale signs of clause nominalization: genitive case-marking on the subject, nominal suffix on the verb, and object case-marking on the entire clause. Nominalized clauses in Turkic languages have all these properties (Lees 1965, Kornfilt 1987, Taylan 1998, Borsley & Kornfilt 2000, Kornfilt 2007, among others). Even though Turkic languages are known to have



Meltem Kelepir. 2021. Matters of size and deficient functional categories in three Turkic languages: Turkish, Turkmen, and Noghay. In Sabine Laszakovits & Zheng Shen (eds.), *The size of things I: Structure building*, 25–44. Berlin: Language Science Press. DOI: 10.5281/zenodo.5524276 

similar nominalization properties in non-main clauses, to my knowledge, no comparative theoretical study has been done so far discussing the morpho-syntactic *differences* in the verbal morphology in complement clauses.

In this chapter, I compare and contrast the morphosyntax of indicative complement clauses in three Turkic languages: Noghay, Turkmen, and Turkish.<sup>1</sup> Turkish and Turkmen belong to the Oghuz branch (Csató & Johanson 1998, Schönig 1998), whereas Noghay belongs to the South Kipchak branch (Csató & Karakoç 1998). The following examples illustrate the clause types that I analyze in this chapter in these three languages:<sup>2</sup>

- (1) Hasan Elif-in gel-diğ-in-i biliyor. Turkish  
Hasan Elif-GEN come-DIK-3POSS-ACC knows  
'Hasan knows that Elif came/is coming/will come.'
- (2) ... telefon-ıñ i:şle-yä:n-nig-in-i  
... telephone-GEN work-IMPF-DIK-3POSS-ACC  
i:şle-me-yä:n-nig-in-i ba:rla-malı. Turkmen<sup>3</sup>  
work-NEG-IMPF-DIK-3POSS-ACC check-NEC  
'... you must check whether the phone is working or not.'
- (3) ... öz borış-ıñız-dı ak yüreg-iñiz ben toltır-ar  
... self debt-2PL.POSS-ACC pure heart-2PL.POSS I pay-IMPF  
eken-iñiz-ge Noghay<sup>4</sup>  
EKEN-2PL.POSS-DAT  
'... (that I hope) that you will pay your own debt with your pure heart ...'  
(Djanbidaeva & Ogurlieva 1995: 125 cited in Karakoç 2007: 354)

Among Turkic languages Turkish is the one that has been studied the most within the generative framework. It has been known that (most) indicative complement

---

<sup>1</sup>This chapter is a condensed and slightly modified version of the manuscript (Kelepir 2013). The reader is invited to consult the manuscript for an extensive discussion of the analyses of complement clauses in Turkish in the literature, for a more detailed explanation of the facts in the languages discussed in this chapter, and for more examples of each argument presented here.

<sup>2</sup>I copied the Noghay and Turkmen examples with the orthography used in the cited sources. The English translations from German and Turkish sources are mine, but I received help from a native speaker of German for two of the German translations. In those cases in which the source does not provide morpheme by morpheme glossing, I have added the morpheme boundaries and glosses based on the translations, discussions in the source, and my knowledge of Turkic morphology. Needless to say, all the errors of interpretation and analysis are mine.

<sup>3</sup>The Turkmen examples are from Clark (1998). Also see Kara (2001) for a grammar of Turkmen.

clauses in Turkish differ from matrix clauses not only in nominalization but also in the absence of most tense-aspect-modality (TAM) markers that are found in matrix clauses. Embedded verbal stems bear what has been traditionally called “nominalizers” that can be preceded by a very small number of verbal inflectional morphemes alongside the verbal root. Typical of such nominalizers in indicative clauses are *-DIK*<sup>5</sup> and *-(y)EcEK*. For reasons of space and expository simplicity, I will use only *-DIK* in my examples and discussion of Turkish in this chapter. (1) above provides an example with it.

One of the central themes in the studies on Turkish complement clauses within the generative framework has been the quest to identify the functional categories that make them up. Two properties of these clauses have made it a real challenge for linguists to come up with a proposal for functional structure: (i) different “nominalizers” seem to be compatible with different TAM properties, so their function doesn’t seem to be solely to nominalize (if at all) (see Keleşir 2015), (ii) even if they’re not analyzed as nominalizers but as some TAM marker, it is hard to tell whether they are the morphological realizations of an inflectional category already present in matrix clauses or of a different one since these morphemes do not straightforwardly correspond to the inflectional morphemes found on matrix verbal stems. Thus, the question whether subordinate clauses in Turkish are *smaller* than CPs and if yes, what kind of a functional category is the complement of a nominalizer head has been a very controversial issue.<sup>6</sup>

The comparative study reported in this chapter shows that smaller complement clauses is not a property all Turkic languages share. In the following, first, I show that Turkish complement clauses are actually smaller than matrix clauses, with additional evidence, and then I argue that the same type of clause in two other Turkic languages, namely Turkmen and Noghay, are not. Furthermore, I argue that even though these three languages differ in the size of their nominalized complement clauses, what they have in common is the deficient nature of the highest inflectional category in the clause. I conclude with the suggestion that nominalization of clauses may not necessarily require smaller clauses or the absence of higher functional *categories* but the absence of higher functional *features*.

<sup>5</sup>Common convention in Turkish linguistics is to represent the consonants and vowels in a morpheme that undergo consonant and vowel harmony in capital letters.

<sup>6</sup>See Kural 1993, 1994, 1998, Kennelly 1996, Göksel 1997, Aygen 2002, Kornfilt 2007, among others.

## 2 Size differences in indicative complement clauses

### 2.1 Background

In order to understand the structure of subordinate clauses in these languages, it is necessary to be familiar with four properties of their morphosyntactic clause structure. These are (i) the different sets of TAM markers and their distribution, (ii) the two types of copular verbs and their distribution, (iii) the two types of negation and their distribution, and (iv) the nature and the distribution of existential predicates. For reasons of space, below I provide facts and examples from Turkish only. However, similar generalizations hold in Noghay and Turkmen as well, as we will see in the following sections.

Turkic languages have rich inflectional verbal morphology and a complex system of morphological combinations, with phonologically contentful as well as phonologically null forms. The TAM suffixes in Turkish are generally categorized into three sets in terms of their ordering on the verbal stem and the combinatorial properties. Set1, closest to the verbal root, contains a couple of modality markers and the negative suffix. Set2 is a large set of aspectual and modality markers. Set3 is relatively small. It contains the past tense and the evidentiality markers, as well as the conditional marker, which I exclude from the discussion in this chapter. The present tense is null. The future marker belongs to Set2. Table 1 provides a visual summary with representative markers.

Table 1: Some of the verbal inflectional markers in matrix clauses

	Set1	Set2	Set3
verb root	negative <i>-mE</i>	necessitative,	past,
+ voice markers	(=Neg1), ability	imperfective, future, possibility	evidential

Mainly, a matrix verbal predicate can be formed in the following ways: the lexical verb can always be optionally inflected with one or more Set1 markers. Once that stem is formed, it has to be inflected with either Set2 or Set3 markers, followed by the agreement markers. In the following, the verb with a Set1 marker (the negative suffix) has combined with a Set3 marker (the past tense marker).

- (4) Ben Elif-i ara-ma-dı-m.  
 I Elif-ACC call-NEG-PAST-1SG  
 ‘I didn’t call Elif.’

In order to express aspectual/modal notions, on the other hand, the lexical verb takes a Set2 marker (in addition to optional Set1 markers). This forms a participial form.<sup>7</sup> There is no overt copula in present tense. The stem is immediately followed by agreement markers in present tense. However, past tense and evidential (Set3) markers occur with a copula. In (5) below, the participial form of the verb carries the necessitative (S2) marker. In this case, the Set3 marker *-di* is preceded by a copular verb *i-*. *i-* has to be inserted when there is a participial form.<sup>8</sup>

- (5) Hasan Elif-i ara-ma-malı i-di.  
 Hasan Elif-ACC call-NEG-NEC COP-PAST  
 ‘Hasan should not have called Elif.’

Thus, the line between Set2 and Set3 in Table 1 above indicates where the copula would be inserted.

Similar to the challenge posed by complement clauses in Turkish, functional structure of matrix clauses also has puzzled generative linguists. This is mainly due to the fact that if one assumes a correspondence between the position of a group of morphemes on the verbal stem and the position of the functional category they realize in morphosyntax, then it is almost impossible to find common inflectional features among the morphemes that occur in the same slot on the verbal stem to propose a position for them in the functional structural hierarchy. For instance, while a number of modality and aspect morphemes occur in the same slot on a verbal stem (Set2), other modality markers occur in different slots (Set1, e.g. ability, and Set3, e.g. evidentiality). These facts have led many researchers to either propose hybrid categories (e.g. Asp/Mod) or no labels at all but just label-less functional categories (e.g. Tense1, Tense2, ... or Zone1, Zone2 etc.).<sup>9</sup> My goal in this chapter is not to propose labels for functional categories.

<sup>7</sup>See Lees 1962, Kornfilt 1996, Göksel 1997, Keleş 2001, Sezer 2001, among others, for a discussion of finite and participle forms.

<sup>8</sup>In Turkish, this copula has three forms: *i-*, its cliticized variant *-y*, which occurs when the copula cliticizes to stems that end with a vowel, and its phonologically null variant, which occurs when it cliticizes to stems that end with a consonant. Even though the clitic variants are more unmarked in modern standard Turkish, for expository reasons, I use only *i-* in all the Turkish examples in this chapter.

<sup>9</sup>For hybrid categories, see Aygen-Tosun (1998), for label-less categories, see Sezer (2001) and Enç (2004).

However, I do assume that there is a correspondence between the morphological ordering of the morphemes and their syntactic positioning. Therefore, I will refer to the label-less functional heads in the syntactic structure as F1, F2, F3. Given that Turkish and the other languages in this chapter are head-final, the ordering  $F1 > F2 > F3$  implies that F3 is the highest functional category in the discussion.

The second piece of information that is crucial in understanding the discussion in the remainder of the chapter is the fact that the languages in this chapter have more than one copular verb with different syntactic distributions. For instance, Turkish, in addition to *i-* ‘be’, as seen in the examples above, has another copula *ol-* ‘be’. *i-* can only be inflected with Set3 markers, not with Set1 or Set2. Thus, I call it the “high copula” (COP). *ol-* ‘be’, on the other hand, can be inflected with any of the markers a lexical verb can. I call it the “low copula” (‘be’). These two copular verbs can also co-occur in a simple clause.

- (6) Ozan burada ol-ma-malı i-di.  
Ozan here be-NEG-NEC COP-PAST  
‘Ozan should not have been here.’

I assume that the low copula is inserted at V and the high copula at F3 to satisfy the requirement for a verbal stem of these categories (Kelepir 2001, Enç 2004).

Turkic languages also have two main negative forms: a negative suffix that attaches to a verbal stem (as shown in Table 1 above) and a free negative form that negates non-verbal forms. The following provide examples from Turkish. The negative suffix, *-mE* (a Set1 marker), is attached to the lexical verb *gel-* in (7a). The non-verbal negative form, *değil*, in (7b), negates the non-verbal predicate ‘at home’ and is followed by the high copula *i-*, which is further followed by past tense and agreement markers.

- (7) a. Hasan gel-me-di.  
Hasan come-NEG-PAST  
‘Hasan didn’t come.’  
b. Ben ev-de değil i-di-m.  
I home-LOC not COP-PAST-1SG  
‘I was not at home.’

It is also useful for the upcoming discussion to label these two negative forms in terms of their height in the structure: the verbal negative suffix *-mE* is the “low negation” whereas the non-verbal free form *değil* is the “high negation” (Kelepir 2001). Double negation structures which show their co-occurrence illustrate this height difference more clearly.

- (8) Hasan bun-u bil-m-iyor değil i-di.  
 Hasan this-ACC know-NEG-IMPF not COP-PAST  
 ‘It was not the case that Hasan didn’t know this.’

Note also that *değil* occurs after and negates the participle in (8). Figure 1 is a rough representation of where I assume these elements may be in the syntactic structure in Turkish.

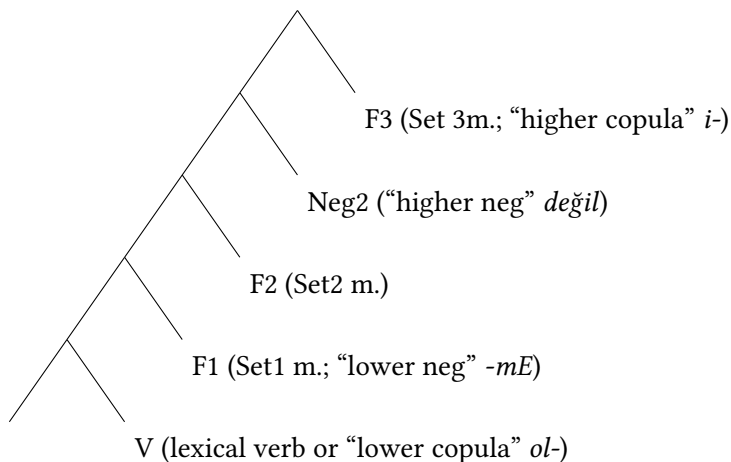


Figure 1: The verb and some functional heads in Turkish

Set1 forms verbal stems whereas Set2 forms participles, hence, non-verbal stems. Set3 markers must attach to verbal stems. This requirement is satisfied by either V (+Set1) or by the high copula. In the latter case, as I mentioned above, the higher copula is inserted at F3 to satisfy the verbal requirement of this category. This is no different from the requirement in English that either the lexical verb gets inflected with, for instance, the past tense marker, *walked*, or in the case of the presence of a participle, the auxiliary/copula does: *was walking*.

Finally, Turkic languages form existential and possessive clauses with special existential predicates. In Turkish, the affirmative form is *var* and the negative form is *yok*. These behave as non-verbal stems, showing combinatorial similarities to nominal and participial forms. For instance, in contrast with lexical verbal roots, they cannot be inflected with any of the Set1 or Set2 markers, as shown in (9a). However, similar to participials (and other non-verbal predicates) but in contrast with verbal stems with only S1 markers, they can be followed by the high copula inflected with Set3 markers, as shown in (9b).

- (9) a. \* Bina-da asansör var-malı.  
building-LOC elevator EXIS-NEC  
'There should be an elevator in the building.'
- b. Bina-da asansör var i-di.  
building-LOC elevator EXIS COP-PAST  
'There was an elevator in the building.'

So, I assume that whatever categories form these existential predicates, as morphosyntactic objects they overlap with participles formed with F2/Set2 and occur below the higher negation in the structure in Figure 1.<sup>10</sup>

With this background in mind, let us now turn to the morphosyntactic properties of indicative complement clauses in Turkish, Noghay, and Turkmen. I start with Turkish and show that these clauses are smaller than main clauses.

## 2.2 Turkish

Embedded verb stems in indicative complement clauses in Turkish differ from the main verb stems in that the number (and the nature) of the inflectional morphemes on the former is much more restricted. Among the three sets of TAM markers I introduced in §2.1, they can only bear Set1 followed by *-DIK*. *-DIK* (alongside with other markers) has been traditionally called a “nominalizer” since it seems to mark the boundary on the stem between the verbal domain (with, for instance, the verbal negative suffix to its left) and the nominal domain (with, for instance, the nominal agreement suffix to its right).

The following is a representative example of the possible morphemes on an embedded verbal stem in indicative complement clauses. The lexical verb *uyu-* ‘sleep’ is (optionally) followed by the negative suffix, then *-DIK*, nominal agreement, and case marking.

- (10) Ozan-ın uyu-ma-dığ-ın-ı biliyorum.  
Ozan-GEN sleep-NEG-DIK-3POSS-ACC I.know  
'I know that Ozan is/was not sleeping.'

---

<sup>10</sup>It is, for instance, possible to create a double negation structure as in the following:

- (i) Bina-da asansör yok değil i-di.  
building-LOC elevator NEG.EXIS not COP-PAST  
'It was not the case that there was no elevator in the building.'



The following illustrate that *-DIK* cannot attach to Set2 (participial stems), as shown in (11a) or to Set3 markers (past tense and evidential markers), as shown in (11b):

- (11) a. \*gel-iyor-duğ-um-u  
           come-IMPF-DIK-1POSS-ACC  
           ‘that I am/was coming’  
       b. \*Ozan-ın Selimiye-de i-di-diğ-in-i  
           Ozan-GEN Selimiye-LOC COP-PAST-DIK-3POSS-ACC  
           ‘that Ozan is/was in Selimiye’

Consequently, embedded clauses tend to be ambiguous with respect to the time, aspect and/or modality of the embedded event, in the absence of corresponding adverbials. At least at first sight, the ungrammaticality of the examples in (11) seems to show that whatever functional categories are realized as Set2 and Set3 markers are missing from embedded clauses. In addition, (11a) shows that *-DIK* cannot attach to a participle, i.e. a non-verbal stem. *-DIK* cannot attach to existential predicates, either.

- (12) \*bu ev-de fare var-diğ-in-i  
       this house-LOC mouse EXIS-DIK-3POSS-ACC  
       ‘that there are/were mice in this house’

Recall that I mentioned in §2.1 that the existential predicates pattern with the participial forms of lexical verbs in their distribution. So, the absence of existential predicates in complement clauses is consistent with the absence of participial forms. I conclude that whatever functional category is responsible for the realization of existential predicates is also absent in these clauses.

Third, they cannot contain the high negation *değil*.

- (13) \*Ozan-ın İstanbul-da değil-diğ-in-i  
       Ozan-GEN İstanbul-LOC not-DIK-3POSS-ACC  
       ‘that Ozan is/was not in İstanbul’

I have been presenting these facts to argue that the verbal domain of these clauses is smaller than that of main clauses. In other words, I propose that the functional categories related to tense, aspect, modality, mood and negation that are higher than the verb phrase (see the structure in Figure 1) must be absent in these subordinate clauses. One might ask whether the absence of participial

forms, existential predicates and *değil* in these clauses may not be due to a morphological requirement of *-DIK* to attach to verbal stems, assuming that these stems may be non-verbal. As plausible an analysis as it sounds, it would not explain the next fact: the fact that these clauses can not contain the high copular verb *i-*, either. As the following example shows, *-DIK* cannot attach to *i-* even though *i-* is a verbal stem (see §2.1 and Figure 1).

- (14) \* Ozan-ın Selimiye-de i-diğ-in-i  
Ozan-GEN Selimiye-LOC COP-DIK-3POSS-ACC  
'that Ozan is/was in Selimiye'

Thus, I conclude that the grammaticality of (10) versus the ungrammaticality of (11–14)<sup>11</sup> point to the fact that Turkish nominalized clauses lack F3 (as well as Neg2) and F2 can only be realized as *-DIK* (and other so-called nominalizers with the same morphosyntactic distribution).<sup>12</sup> I argue in detail in Kelepir (2013) that embedded F2 lacks morphosyntactic aspect and modality features, and *-DIK* as an underspecified, default morpheme is inserted at this category. In short, the highest functional (inflectional) category in these embedded clauses below the head that is responsible for nominalization is F2 and it is deficient in terms of aspect/modality morphosyntactic features.

Having seen that nominalized complement clauses have a very small verbal domain in Turkish, the immediate question that arises is whether the functional category that is responsible for nominalization of the clause can only select for a small verbal domain in Turkic languages in general. A careful analysis of Noghay and Turkmen shows that this is not the case. I start with Noghay.

<sup>11</sup>Sağ (2013) reports that these structures are grammatical in Denizli dialect of Turkish. Moreover, *-DIK* attaching to the high copula *i-*, as in (14), was possible in Ottoman Turkish (Kerslake 1988: 195) and the form *idüğü* remains in an idiomatic, frozen form in modern Turkish (Banguoğlu 1990, Kelepir 2013).

<sup>12</sup>The grammatical counterparts of the ungrammatical examples in (11–14) require the low copula *ol-*, which *-DIK* can attach to, see (i) below. *ol-* behaves like any other lexical verb morphosyntactically, so I assume that it is inserted at V and takes the non-verbal predicates as its complement.

- (i) a. geli-iyor ol-duğ-um-u (compare with (11a))  
b. bu ev-de fare ol-duğ-un-u (compare with (12))  
c. Ozan-ın İstanbul-da ol-ma-dığ-in-ı (compare with (13))  
d. Ozan-ın Selimiye-de ol-duğ-un-u (compare with (14))

See also Kerslake (1988) and Göksel (2001) for a detailed analysis of the functions of *ol-* in matrix and embedded clauses.

### 2.3 Noghay

Noghay is similar to Turkish in that the main clause verbal predicate may consist of a finite verb or a participle plus the high copula. The cognate of the high copula *i-* of Turkish in Noghay is *e-*. In her description of Noghay grammar, Karakoç (2001) reports that the copula *e-* has three inflected forms: the past form *edi*, the indirective-modal copula form *eken*, and the conditional form *ese*. Karakoç (2001) also notes that *e-* is not a regular lexical verbal root, i.e. it cannot be used as a full verb, it can only be inflected with the morphemes mentioned above. Recall that these are similar to the only morphemes (Set3 morphemes) that the Turkish high copula *i-* can carry. In the discussion of Noghay, my focus will be the form *eken*. Karakoç (2001) observes that *eken* expresses the notions “evidential”, “inferential” and/or “indirective”, among others. Henceforth, I will use the term “evidential” as a cover term for all these related meanings.

As expected from a copular form, *eken* can occur with non-verbal predicates, as in (15a) and with participial forms, as in (15b).

(15) Noghay

- a. Ali eginši e-ken.  
Ali farmer COP-EVID  
‘Apparently, Ali was/is a farmer.’ (Karakoç 2001: 23)
- b. ... sen bir âlemet bol-ıan e-ken-sin  
you a strange be-PERF COP-EVID-2SG  
‘... (as I find out) you have become strange’ (Kazakov 1983: 33, cited in Karakoç 2001: 25)

I propose that similar to the main clause structure in Turkish, the evidential marker *-ken* in Noghay is a Set3 marker (recall Table 1) and thus, is inserted into F3 (recall the structure in Figure 1). Similar to the Turkish high copula *i-*, Noghay high copula *e-* is inserted into F3 to satisfy the verbal requirement of this category since neither the nominal predicate (15a) nor the participle (15b) can do it.

However, Noghay nominalized indicative complement clauses differ from those in Turkish in a number of respects. First of all, there is no morpheme that may be easily considered a “nominalizer” such as the morphologically more opaque form *-DIK* in Turkish. Instead, the high copula form *eken* carries nominal agreement (possessive) and case morphology.

- (16) ... bayının            anası    eken-in            de    anlaydı.  
her.husband's mother EKEN-3POSS.ACC also realizes  
'(The woman) ... realizes that she is her husband's mother.' (Karakoç 2007: 343)

The presence of a high copula in the subordinate clause raises the question whether these clauses are bigger than their counterparts in Turkish. As a matter of fact, they are. They can contain participles, (3) and (17), existential/possessive predicates *bar/yoq*, (18), and the non-verbal negator *tuwıl*, (19).

- (17) Kılıplı karttın    sözi kim-ge    tiy-edi            eken-in            B.  
sneaky old.man's word who-DAT touch-IMPF EKEN-3POSS.ACC B.  
anladı.  
understood  
'Baymurza understood whom the sneaky old man's words targeted (and ...).' (Djanbidaeva & Ogurlieva 1995: 126, cited in Karakoç 2007: 353)
- (18) Kim bileđi bu    oyırsızdın            yüreg-in-de            ne    bar  
who knows this good.for.nothing's heart-3POSS-LOC what EXIS  
eken-in.  
EKEN-3POSS.ACC  
'Who knows what is in this good-for-nothing's heart.' (Djanbidaeva & Ogurlieva 1995: 55, cited in Karakoç 2007: 344)
- (19) İdris /.../ kelininiñ quwnaq tuwıl eken-in            körip  
İdris    his.bride's good    not    EKEN-3POSS.ACC saw.and  
'... İdris saw that his daughter-in-law was not well ...' (Kapaev 1962: 159 cited in Karakoç 2001: 33)

In short, in this section I have shown that the nominalized indicative complement clauses in Noghay have a verbal functional structure similar to that of main clauses, in contrast to Turkish.

What is striking is that in these clauses *eken* does not function as a modal marker expressing evidentiality, as it does in main clauses. Karakoç (2001) reports that it does not express evidentiality but functions only as a static copula. In other words, it is semantically vacuous. Thus, I propose that even though *eken* with the high copula resides in F3 (see Figure 1), embedded F3 lacks the morphosyntactic features it may bear in matrix clauses. Hence, the lack of evidential interpretation. Thus, I suggest that similar to *-DIK* in Turkish, *-ken* in *eken* is an underspecified morpheme (Keleşir 2015). That is why it is inserted into a deficient F3. I return to this point in §3.

## 2.4 Turkmen

Turkmen, which belongs to the East Oghuz group (Schönig 1998, Johanson 1998), is a language genetically closer to Turkish than Noghay, and even though Turkmen and Turkish seem to be very similar on the surface, there is a striking difference between the two languages in the morphosyntax of complement clauses. Consider the following example from Turkmen where the embedded verbal stem bears the marker *-DIK*:

- (20) Turkmen  
 ... nä:me tölö-mölü-düg-ü    şol ta:yda aydyla:r  
 ... what pay-NEC-DIK-3POSS there    is.said  
 ‘... it is said there what you have to pay’

(20) shows that in contrast with Turkish where the complement clause cannot have a verbal stem containing a Set2 marker, Turkmen allows this. The verbal stem contains a Set2 marker, one of the allomorphs of the necessitative marker *-mEll*. Thus, *-DIK* seems to attach to a participial form. The examples in Clark (1998: 480–483) also include other participial markers such as the future participle, the present participle marker, and the past participle marker that *-DIK* attaches to.

Does this show a difference between the morphological requirements of the *-DIK* markers in Turkish and Turkmen? Namely, can the Turkmen *-DIK* attach to a participial marker where the Turkish *-DIK* cannot? Or is the difference syntactic?

A closer look suggests that it is syntactic. Recall that Turkish (and Noghay) have the high copular verbs. In the Turkish examples we saw earlier, the high copula had the form *i-*. In Footnote 8 I noted that it also has two clitic variants: *-y* and a phonologically null form. When we analyze the Turkmen matrix clause examples where the copula is expected to occur, we see that it is never realized with phonological content. Even in environments where the Turkish copula would either be *i-* or the clitic *-y*, it is phonologically null. Contrast the Turkish and Turkmen examples in (21). In both, the stem ends with a vowel, *u*. Turkish copula is in the form of clitic *-y* (or it can be *i-*, but it cannot be null), whereas the Turkmen copula is null.

- |   |  |
|---|--|
| (21) a. Dolu-y-du.<br>full-COP-PAST<br>‘It was full.’ (Turkish) | b. Do:lu-Ø-di.<br>full-COP-PAST<br>‘It was full.’ (Clark 1998: 239)<br>(Turkmen) |
|---|--|

I would like to propose that the reason why *-DIK* attaches to Set2 markers, as in (20) is that the syntactic position of *-DIK* is the same as the position of *-ken* in Noghay. It realizes a high functional (inflectional) category, F3, which requires a verbal predicate. In the absence of a verbal stem, this high functional category hosts the high copula, which is phonologically null. However, neither the phonologically null copula nor the suffix *-DIK* can occur on their own, they have to attach to a stem to their left. So, (20) should actually be represented as (22) below.

- (22) ... nā:me tölö-mölü-Ø-düg-ü şol ta:yda aydılyar  
 what pay-NEC-COP-DIK-3POSS there is.said  
 ‘... it is said there what you have to pay’

If *DIK* really attaches to a null copula, an immediate prediction is that it should also attach to a non-verbal predicate. This is borne out.

- (23) Turkmen: non-verbal predicate  
 O-nurj a:ga-m-Ø-diğ-in-i derrew tanadım  
 he-GEN older.brother-1POSS-COP-DIK-3POSS-ACC immediately I.recognized  
 ‘I recognized immediately that he was my older brother.’

The possibility of the occurrence of participial markers plus *-DIK* in Turkmen implies that complement clauses in Turkmen contain all the three functional categories, in contrast with Turkish, but similar to Noghay.<sup>13</sup>

If that is the case, then Turkmen complement clauses should be able to contain the high negation (the non-verbal negator) and existential predicates. This is, in fact, the case. In the examples below the high negation is *däl*, (24), and the existential predicate is *bar*, (25).

- (24) O-nun   gowı mugallıma däl-Ø-dig-in-i                      eşitdim.  
she-GEN good teacher     not-COP-DIK-3POSS-ACC I.heard  
'I heard that she is not a good teacher.'

<sup>13</sup> An anonymous reviewer suggests that *-DIK* in Turkmen must be etymologically related to *-LIK*, which in many Kipchak and Turkic languages follows the non-finite clause head, and that Turkmen must have borrowed it from neighboring Kipchak languages (Asarina (2011) analyzes *-LIK* as a complementizer in Uyghur whereas Ótött-Kovács (2018) analyzes it as a nominalizer in Kazakh). However, I don't think Turkmen *-DIK* is an allomorph of *-LIK*. Clark (1998: 480–483) describes *-DIK* in Turkmen as a particle separate from *-LIK* and reports that the use of each morpheme causes a different interpretation. He states that while *-LIK* emphasizes the nominal character of the object, *-DIK* emphasizes its verbal character. He translates those with *-DIK* as *that*-clauses whereas those with *-LIK* as gerunds, e.g. 'never forget about a stick's having two points' or 'realized about my having made a mistake'.

- (25) Ol meniň pulumyň bar-Ø-dig-in-1 bilyär.  
s/he my money EXIS-COP-DIK-3POSS-ACC knows  
‘She knows that I have money.’

Notice that even though the embedded sentence in (23) is translated with past tense ‘was my brother’, there is no past tense marker (*-DI*) on the embedded predicate. Contrast this with the past tense marker on the adjectival predicate in the matrix clause in (21b). A comparison of the embedded predicates in (23–25) shows that even though in all these there is no embedded tense marking, (23) is interpreted with past tense whereas those in (22), (24), and (25) with present tense. This is in fact reminiscent of the ambiguous tense (and aspect) interpretation in embedded clauses in Turkish mentioned in §2.2. In both languages the tense interpretation seems to rely on the tense of the main verb and context. While Turkmen embedded clauses with *-DIK* are ambiguous only with respect to tense, Turkish embedded clauses with *-DIK* are ambiguous with respect to both aspect (F2 category) and tense (F3 category). Thus, for Turkmen, I conclude that even though it contains both F2 and F3 categories, F3 is deficient in terms of morphosyntactic features, similar to Noghay.

### 3 Conclusion and implications

We have seen that even though in all of these three languages complement clauses are nominalized, they differ in the size of the clausal (verbal) domain below the nominal domain: the clausal domain is smaller, with fewer functional (inflectional) categories, than the main clauses in Turkish, whereas it is *almost* as big as the main clauses in Noghay and Turkmen. Noghay and Turkmen contain functional categories expressing aspectual and modal differences, forming existential and possessive predicates, the high negation, and the high copula. What this implies is that nominalization does not necessarily require a smaller clause, at least in Turkic languages. One way of accounting for the difference between these languages could be proposing that each nominalizing functional head in each language has a different selectional requirement, resulting in complement clauses with different sizes.

However, the nominalizing head seems to still put a requirement on the head of its complement: that it should be deficient, devoid of any morphosyntactic tense and evidentiality features. The most straightforward evidence for this comes from Noghay data. Recall that Noghay is reported to have two finite high copula forms: *eken* and *edi*. In main clauses, *eken* expresses evidentiality and *edi*

expresses past tense. Even though Noghay embedded clauses appear to be as big as the main clauses, there are crucial differences: first of all, *edi* cannot occur in the embedded clause. Second, as mentioned in §2.3, even though *eken* does occur in embedded clauses, it does not express evidentiality, in fact, it is devoid of any meaning (Karakoç 2007). So, even though the presence of the high copula *e-* shows that the highest functional (inflectional) category is present in the embedded clause, the impossibility of the “high” marker past tense *-di*, and the meaninglessness of the other “high” marker *-ken* point to the conclusion that highest functional category (F3) in embedded clauses lacks the morphosyntactic features it may bear in matrix clauses. I have argued in Keleşir (2015) that the reason why *-ken* is inserted into F3 is that it is the allomorph of an underspecified marker, *-GAn*, as opposed to the F3 marker *-DI* in *edi*, which is specified for past tense.<sup>14</sup>

Similarly, the highest functional categories, in Turkmen and Turkish, F3 and F2, respectively, seem to be deficient with respect to the morphosyntactic features they carry in main clauses. If this is correct, then we observe a dissociation of morphosyntactic features of heads from their semantic features. In all the three languages analyzed here, the embedded clauses can express aspectual and tense properties independent from the matrix clause, implying that the related operators are actually present. The dissociations between inflectional morphology from the semantics of related inflectional notions (e.g. tense) is familiar from works on sequence-of-tense phenomena and discussions on tensed vs. tenseless infinitives (Stowell 1982, Wurmbrand 2014, see also Enç 1987 and Ogiwara 1996). In the particular case of Noghay, we see that even when the evidential *morphology* is present in the embedded clause, the *semantics* of evidentiality is absent. Namely, the evidential marker is semantically vacuous in embedded contexts.<sup>15</sup>

Throughout the chapter, I have refrained from labeling the functional categories that are absent or present in embedded clauses. As I mentioned in §1, this

---

<sup>14</sup> An anonymous reviewer asks whether an alternative theory could be proposed: that *eken* has been reanalyzed and is now its own lexical entry with its own syntax and its meaning, and that this would explain why it differs from *edi* and it has an unexpected meaning. *Eken* may have been reanalyzed and grammaticalized as a monomorphemic element. In fact, Karakoç (2001, 2007) treat it that way. However, it does not have its own syntax since its position and what complements it can take are not different from *eken* in main clauses. The difference is in the interpretation. The analysis I am arguing for here is meant to raise the question why among the two copular forms, *eken* and *edi*, it is *eken* that is used in nominalized clauses, or if we adopt the reviewer’s alternative theory, why it is *eken* that got reanalyzed. My answer is because perhaps it contains an underspecified morpheme, which functions as a default F3 marker, whereas *edi* does not.

<sup>15</sup> See Aikhenvald (2004) and Schenner & Sauerland (2007) on the question whether evidentials can be embedded.



is mainly due to the reason that the set of morphemes that occupy the same positions in the embedded verbal template do not seem to have a common inflectional feature (i.e. aspect, modality, tense, mood etc.). However, I have also refrained from even claiming whether or not “big” clauses in Noghay and Turkmen are CPs or not. Since, for instance, I argue that the high copulas *e-* in Noghay and  $\emptyset$  in Turkmen are inserted at the “highest inflectional category” (F3), one might wonder whether these clauses are full CPs as in matrix clauses. Note that recent work by Susi Wurmbrand with Magdalena Lohninger (Wurmbrand & Lohninger 2019) analyzes nominalized complement clauses in Buryat (Bondarenko 2018), a Mongolian language spoken in the Russian Federation, in relation to their proposal for a *universal implicational complementation hierarchy* (ICH), and claims that these clauses do not display CP-hood characteristics and thus must be smaller than CPs. What I argued for in this chapter and what Wurmbrand & Lohninger (2019) propose do not necessarily contradict each other. Further research on Turkish, Noghay and Turkmen (and possibly other languages with nominalized complement clauses) may point to a more fine-grained layering of the “highest functional categories” and/or of the “highest” morphosyntactic features in nominalized embedded clauses.

## Acknowledgments

I thank the two anonymous reviewers for their invaluable feedback and the editors for their support and patience. All errors are mine.

## Abbreviations

ACC	accusative	IMPF	imperfective	PL	plural
COP	copula	LOC	locative	POSS	possessive
DAT	dative	NEC	necessitative	SG	singular
EVID	evidential	NEG	negative		
EXIS	existential	PAST	past		
GEN	genitive	PERF	perfective		

## References

Aikhenvald, Alexandra Yurievna. 2004. *Evidentiality*. Oxford: Oxford University Press.

- Asarina, Alevtina. 2011. *Case in Uyghur and beyond*. Cambridge: Massachusetts Institute of Technology. (Doctoral dissertation).
- Aygen, Gülşat. 2002. *Finiteness, case and clausal architecture*. Cambridge: Harvard. (Doctoral dissertation).
- Aygen-Tosun, Gülşat. 1998. *The SPLIT INFL hypothesis in Turkish*. Istanbul: Bogazici University. (M.A.).
- Banguoğlu, Tahsin. 1990. *Türkçe'nin grameri [Turkish grammar]*. Ankara: Türk Tarih Kurumu Basımevi.
- Bondarenko, Tatiana. 2018. Strategiji nominalizacii v barguzinskom dialekte buryatskogo yazyka (Nominalization strategies in Barguzin Buryat). *Ural-Altaic Studies* 4(31). 95–111.
- Borsley, Robert & Jaklin Kornfilt. 2000. Mixed extended projections. In Robert Borsley (ed.), *The nature and function of syntactic categories* (Syntax and Semantics 32), 101–131. San Diego: Academic Press.
- Clark, Larry. 1998. *Turkmen reference grammar* (Turcologica 34). Wiesbaden: Harrasowitz Verlag.
- Csató, Éva Á. & Lars Johanson. 1998. Turkish. In Lars Johanson & Éva Á. Csató (eds.), *The Turkic languages*, 203–235. New York: Routledge.
- Csató, Éva Á. & Birsel Karakoç. 1998. Noghay. In Lars Johanson & Éva Á. Csató (eds.), *The Turkic languages*, 333–343. New York: Routledge.
- Djanbidaeva, Djenetchan K. & E. S. Ogurlieva (eds.). 1995. *Nogay literaturası. Çerkessk: Karaşay-Şerkeş kitap baspası*.
- Enç, Mürvet. 1987. Anchoring conditions for tense. *Linguistic Inquiry* 18(4). 633–657.
- Enç, Mürvet. 2004. Copulas and functional categories in Turkish. In Aniko Csirmaz, Youngjoo Lee & Mary Ann Walter (eds.), *Proceedings of the 1st Workshop of Altaic formal linguistics (WAFL 1)* (MIT Working Papers in Linguistics), 208–226. Cambridge: MIT Press.
- Givón, T. 2009. *The genesis of syntactic complexity: Diachrony, ontogeny, neuro-cognition, evolution*. Amsterdam: John Benjamins.
- Göksel, Aslı. 1997. Word length. In Geert Booij, Angela Ralli & Sergio Scalise (eds.), *Proceedings of the 1st Mediterranean conference of morphology*, 190–200. Patras: University of Patras.
- Göksel, Aslı. 2001. The auxiliary verb at the morphology-syntax interface. In Eser E. Taylan (ed.), *The verb in Turkish*, 151–182. Amsterdam: John Benjamins.
- Johanson, Lars. 1998. The history of Turkic. In Lars Johanson & Éva Á. Csató (eds.), *The Turkic languages*. New York: Routledge.
- Kapaev, İsa. 1962. *Eski üjdin soñy*. Çerkessk: Stavropol' kniga baspası.
- Kara, Mehmet. 2001. *Türkmençe*. Ankara: Akçağ Yayınları.

- Karakoç, Birsal. 2001. *Das finite Verbalsystem im Noghaischen*. Mainz: Johannes Gutenberg University. (Doctoral dissertation).
- Karakoç, Birsal. 2007. Nogayca ve Türkiye Türkçesi'nde tümleç yan cümlelerinde kullanılan yüklemleştircilerin bir karşılaştırması. In László Károly (ed.), *Turcology in Turkey: Selected papers*, 337–359. Szeged: SZTE Altajisztikai Tanszék & MTA-SZTE Turkológiai Kutatócsoport.
- Kazakov, Valerij. 1983. *Togyz kaptal: Povest'ler men xabarlar*. Çerkesk: Stavropol' kniga baspası.
- Kelepir, Meltem. 2001. *Topics in Turkish syntax: Clausal structure and scope*. Cambridge: Massachusetts Institute of Technology. (Doctoral dissertation).
- Kelepir, Meltem. 2013. The presence/absence of tense and evidentiality in nominalized complement clauses: Some insights from three Turkic languages. Istanbul.
- Kelepir, Meltem. 2015. Nominalization morphemes are underspecified participial markers. In Deniz Zeyrek, Çiğdem S. Şimşek, Ufuk Ataş & Jochen Rehbein (eds.), *Ankara papers in Turkish & Turkic linguistics*, vol. 103 (Turcologica), 110–120. Wiesbaden: Harrasowitz Verlag.
- Kennelly, Sarah D. 1996. Turkish subordination: [–Tense, –CP, +Case]. In Ahmet Konrot (ed.), *Proceedings of the 6th international conference on Turkish linguistics* (Modern studies in Turkish linguistics), 55–75. Eskişehir: Anadolu University.
- Kerslake, Celia. 1988. Semantic differentiation in the copular system of Modern Turkish. In Sabri Koç (ed.), *Studies on Turkish linguistics*, 147–185. Ankara: Middle East Technical University.
- Kornfilt, Jaklin. 1987. Turkish and the Turkic languages. In Bernard Comrie (ed.), *The world's major languages*, 619–644. New York: Oxford University Press.
- Kornfilt, Jaklin. 1996. On some copular clitics in Turkish. In Artemis Alexiadou, Nanna Fuhrhop, Paul Law & Sylvia Lohken (eds.), *ZAS papers in linguistics*, 96–114. Berlin: ZAS.
- Kornfilt, Jaklin. 2007. Verbal and nominalized finite clauses in Turkish. In Irina Nikolaeva (ed.), *Finiteness: Theoretical and empirical foundations*, 305–332. Oxford: Oxford University Press.
- Kural, Murat. 1993. V-(to)-(I-to)-C in Turkish. In Filippo Beghelli & Murat Kural (eds.), *UCLA occasional papers in linguistics*, vol. 11, 1–37. Los Angeles: UCLA Department of Linguistics.
- Kural, Murat. 1994. Infl and agreement in Turkish subordinate clauses. In Donka Farkas, Erin Duncan & Philip Spaelti (eds.), *The proceedings of the 12th West Coast Conference on Formal Linguistics*, 339–354. Cambridge: Cambridge University Press.

- Kural, Murat. 1998. Subordinate Infs and Comp in Turkish. In Lars Johanson, Éva Á. Csató, Vanessa Locke, Astrid Menz & Dorothea Winterling (eds.), *The Mainz meeting: Proceedings of the 7th International Conference on Turkish Linguistics* (Turcologica 33), 404–421. Wiesbaden: Harrasowitz Verlag.
- Lees, Robert B. 1962. A compact analysis for the Turkish personal morphemes. In Nicholas Poppe (ed.), *American studies in Altaic linguistics*, vol. 13 (Indiana University Uralic and Altaic Series), 141–176. The Hague: Mouton.
- Lees, Robert B. 1965. Turkish nominalization and a problem of ellipsis. *Foundation of Language* 1(2). 112–121.
- Ogihara, Toshiyuki. 1996. *Tense, attitudes, and scope*. Dordrecht: Kluwer.
- Ötö-Kovács, Eszter. 2018. Kazakh non-finite clauses followed by *-LIQ* as a case in favor of the clause-internal nominalization hypothesis. In Jaklin Kornfilt & Yilmaz Koylu (eds.), *Thematic volume: Papers in Turkish and Turkic linguistics*, vol. 18 (IULC Working Papers 2). Bloomington: Indiana University.
- Sağ, Yağmur. 2013. Copula in Turkish. In Umut Özge (ed.), *Proceedings of the 8th Workshop on Altaic Formal Linguistics (WAFL)*, vol. 67, 293–298. Cambridge: MIT Working Papers in Linguistics (MITWPL).
- Schenner, Mathias & Uli Sauerland. 2007. Shifting evidentials in Bulgarian. In Estela Puig-Waldmüller (ed.), *Proceedings of Sinn und Bedeutung 11*, 495–509. Barcelona: Universitat Pompeu Fabra.
- Schönig, Claus. 1998. Turkmen. In Lars Johanson & Éva Á. Csató (eds.), *The Turkic languages*, 261–272. New York: Routledge.
- Sezer, Engin. 2001. Finite inflection in Turkish. In Eser E. Taylan (ed.), *The verb in Turkish*, 1–45. Amsterdam: John Benjamins.
- Stowell, Timothy. 1982. The tense of infinitives. *Linguistic Inquiry* 13(3). 561–570.
- Taylan, Eser E. 1998. What determines the choice of nominalizer in Turkish? In Bernard Caron (ed.), *Proceedings of the 16th International Congress of Linguists*, paper no.220 (CD Rom). Netherlands: Emerald Group Publishing Limited.
- Wurmbrand, Susi. 2014. Tense and aspect in English infinitives. *Linguistic Inquiry* 45. 403–447.
- Wurmbrand, Susi & Magdalena Lohninger. 2019. An implicational universal in complementation: Theoretical insights and empirical progress. In Jutta M. Hartmann & Angelika Wöllstein (eds.), *Propositional arguments in cross-linguistic research: Theoretical and empirical issues*. Berlin: Mouton de Gruyter.