

# Remarks on Compounding in Turkish

Aslı Göksel and Belma Haznedar  
Boğaziçi University

1. Introduction
2. Stress
3. Headedness
4. Compounding by reduplication
5. The N/A distinction and its implications for the SUB/ATT distinction
6. Linking elements
7. Higher order compounds
8. Recursivity
9. Inflection

## 1. Introduction

This paper presents a general overview of the structure of compounds in Turkish and should be read in conjunction with the Table of Turkish Compounds. After a brief description of the main properties of compounds, the layout of the rest of the paper is as follows: we discuss the stress pattern of compounds in section 2 as a preamble to issues concerning the status of heads in section 3. Section 4 is on compounding through doubling as a derivational process. In section 5 we discuss the issues concerning the differences between adjectives and nouns in Turkish since some of the distinction between attributive and subordinate compounds hinges partly on this factor. Section 6 is about linking elements (LE), suffixes that are used for compounding purposes in Turkish. In section 7 we present a brief description of clausal (higher order) compounds, followed by recursivity in compounds in section 8 and inflection in section 9.<sup>1</sup>

### 1.1 Basic categories

As in many other languages, compounding in Turkish is a productive and regular process.<sup>2</sup> While Turkish has compounds in the three major nominal categories N, A and ADV, the vast majority of compounds belong to the syntactic category of nouns and are made up of two nouns and a linking element  $[[N+N]+LE]_N$ :

---

<sup>1</sup> This research was supported by the Boğaziçi University Research Fund, Project # 06B401. The abbreviations we use in this paper are as follows: A: adjective, ABL: ablative, ACC: accusative, ADV: adverb, ATT: attributive, DAT: dative, INFL: inflection, N: noun, P: past, POSS: possessive, PRTC: participle, OPT: optative, PL: plural, SG: singular, SUB: subordinate, V: verb, LE: linking element. Capital letters in the suffixes refer to the variable phonological segments.

<sup>2</sup>For descriptions and analyses of compounding in Turkish, the reader is referred to Dede (1978) and van Schaik (2002) which are the most detailed studies on the topic.

- (1) a. *para çanta-sı* (money bag) ‘wallet’  $[[N+N]+LE]_N$   
 b. *göz nur-u* (eye light) ‘toil’  $[[N+N]+LE]_N$   
 c. *hamuriş-i* (dough work) ‘pasta’  $[[N+N]+LE]_N$   
 d. *köpek diş-i* (dog tooth) ‘canine tooth’  $[[N+N]+LE]_N$

These compounds consist of two simple or complex nouns where the second constituent is marked with *-(s)I(n)*. With only a few types of (putative) exceptions (see 6.1), the non-head of these compounds are nouns.

N+N compounds without a linking element are also common but restricted to certain types. While N+N compounds with a linking element are usually of the subordinating type those without a linking element are attributive (see Bisetto and Scalise 2006). These include examples where the first noun is in a subordination relation with the head (2a), or where it is used attributively (2b). The latter are cases where the compound denotes forms of address (2c), or names of districts, companies etc. For a discussion of N+N compounds without linking elements, see Gökdayı (2007). For the optionality of the linking element, see section 6.2.

- (2) a. *babaanne* (father mother) ‘grandmother’  $[N+N]_N$   
 b. *demir kapı* (iron door) ‘iron door’  $[N+N]_N$   
 c. *Hasan Bey* (*Hasan* + a form of address)  $[N+N]_N$   
 d. *Topkapı* (cannon gate) a place name  $[N+N]_N$

$[A+N]_N$  type of compounding is also productive:

- (3) a. *darboğaz* (narrow strait) ‘bottleneck’  $[A+N]_N$   
 b. *sarıkanat* (yellow wing) ‘blue fish’  $[A+N]_N$   
 c. *dar açı* (narrow angle) ‘acute angle’  $[A+N]_N$

Lastly, compounds which are nouns can also be formed by the combination of two verbal roots:

- (4) a. *gelgit* (come go) ‘tide’  $[V+V]_N$   
 b. *çekyat* (pull lie.down) ‘sofa bed’  $[V+V]_N$

Compounds which are adverbs (5) and adjectives (6) are discussed in sections 4.1 and 5.1 respectively.

- (5) a. *yaz kış* (summer winter) ‘continuously’  $[N+N]_{ADV}$   
 b. *yavaş yavaş* (slow slow) ‘slowly’  $[A+A]_{ADV}$
- (6) a. *elma yanak-lı* (apple cheek-A) ‘rosy cheeked’  $[N+A]_A$

b. *kan kırmızı* (blood red) ‘cadmium red’ [N+A]<sub>A</sub>

## 1.2 (Non-)Headedness

Turkish endocentric compounds are mostly right-headed:

- (7) a. *baş örtüsü* (head cover) ‘kerchief’  
b. *deniz otobüsü* (sea bus) ‘(public) hovercraft’

It should be noted, however, that apart from left-headed Arabic compounds which are obsolescent (8a), there are two types of native compounds for which a ‘left-headed’ analysis seems more plausible (8b-c).

- (8) a. *kabil-i tahammül* (able-LE tolerate), ‘tolerable’  
b. *dut kuru-su* (mulberry dry-LE) ‘dried mulberries’  
c. *tirşe reng-i* (turquoise colour-LE) ‘the colour turquoise’

Headedness will be discussed in more detail in section 3. Exocentric compounds are discussed in sections 2.3.1 and 3.2.

## 1.3 The relationship between the constituents of compounds

With respect to the relationship between the constituents inside a compound, we follow the three way distinction discussed in Bisetto and Scalise (2006), which also underlies the classification in the Table. According to this analysis the constituents of a compound can be in a subordinating, in an attributive/appositive or in a coordinated relation. These are henceforth abbreviated as SUB, ATT or ATAP and COOR respectively. All types occur in Turkish.<sup>3</sup>

## 2. Stress in compounds

### 2.1 Regular stress

Compounds in Turkish are stressed on the stressable syllable of the non-head, which is usually the final syllable of the first item:

- (9) a. *kabák* çekirdeğ-i  
pumpkin seed-LE  
‘pumpkin seed’  
b. *kurú* kayısı  
dry apricot  
‘dried apricots’

<sup>3</sup> See Dede (1987) and van Schaaik (2002) for the types of relation between the two constituents of compounds in Turkish.

c. *Á*nkara armud-u  
 Ankara pear-LE  
 ‘(type of) pear’

d. ev-lí bark-lı  
 house-A sW-A  
 ‘married’

We shall refer to such compounds as *regularly stressed compounds*.

A small number of compounds is stressed on the constituent on the right. One such type is the group of exocentric compounds below. Stressing the constituent on the right is a general rule that applies to these. We shall discuss such compounds in more detail in section (3.1-2):

(10) göz-ü pék  
 eye-POSS sturdy  
 ‘courageous’ (lit. her/his eye (is) sturdy)

We refer to compounds which are not regularly stressed as *finally stressed compounds*.

## 2.2. The prosodic structure of compounds

The stress pattern of compounds has been subject to several analyses investigating the structure of stress in Turkish. Earlier approaches take the view that as a result of reanalysis, stress is assigned to the leftmost element among two primary stressed ones within a domain. According to these analyses the stress of the items on the right is reduced, resulting in what we call here regularly stressed compounds (Lees 1961). Among the more recent approaches, Kabak and Revithiadou (2006), (2007) suggest that the stress pattern of compounds, together with the stress pattern found in constituents containing clitics, identifies a particular prosodic level, which they call the Clitic Group. Within this non-derivational approach, the Clitic Group is larger than the ‘prosodic word’ but smaller than the ‘prosodic phrase’. The leftmost constituent of the Clitic Group is the prosodic word with stress on its left edge and outside of this domain and to the right of the phonological word is either a group of clitics or the unstressed item of a compound. A different, but also non-derivational approach is put forward by Charette, Göksel and Şener 2007. On this view, the stress pattern of compounds is the manifestation of a binary structure, the Templatic Phrase, which is the structure that underlies a stress-head relation. Within this binary structure the left branch hosts stress and the right branch hosts the head. Such a structure encompasses various types of construction: syntactic phrases, partially and fully reduplicated stems and, of course, compounds.

## 2.3 Finally stressed compounds

While all the approaches above account for regularly stressed compounds as in (9), the stress pattern of type of compound in (10), together with other finally stressed compounds, remains mostly unexplained. Finally stressed compounds do not constitute a uniform group except that right-headed finally stressed compounds are extremely rare. The following belong to the exceptional group of finally stressed compounds which can

be right-headed (11a), or SUB and exocentric where the subordinating item is on the right:

- (11) a. er+bâş  
           soldier head  
           ‘noncommissioned officer’
- b. merkez+kâç  
           center escape  
           ‘centrifugal’

For derivational approaches, the leftmost item in finally stressed compounds would probably lack primary stress at the relevant level. For Kabak and Vogel (2001) and Kabak and Revithiadou (2006), (2007) these compounds would probably constitute a phonological word and not be part of the Clitic Group. For Charette, Göksel and Şener (2007) these are an exception to the Templatic Phrase, for although they have a stressed non-head, this non-head is on the right. Hence, the construct Templatic Phrase as described in Charette, Göksel and Şener (2007) which is based on the relationship between stress and, crucially, a single head would automatically exclude such structures. So all approaches would, at this point, fail to explain finally stressed constructions containing more than a single stem, which is a crucial part of the description of compounds.

Below, we give a description of the types of finally stressed compound in Turkish. These groups are:

- (i) (COOR) exocentric finally stressed compounds (2.3.1)
- (ii) (SUB) exocentric finally stressed compounds (2.3.2)
- (iii) double-headed finally stressed compounds (2.3.3)
- (iii) finally stressed compounds borrowed from Persian (2.3.4)
- (iv) finally stressed compounds borrowed from Arabic (see section 3.2)

### 2.3.1 COOR exocentric finally stressed compounds

Most of these contain verbs in both constituents and these verbs may be bare, inflected with tense or in the participial form:

- (12) a. gel                   gít  
           come               go  
           ‘tide’
- b. kap-tı               kaç-tı  
           grab-PAST       run.away-PAST  
           ‘public car’
- e. bilgi say-ár  
           information compute-PRT  
           ‘computer’

However not all finally stressed compounds which are exocentric contain verbs (e.g. (13a) and (13b)), nor are all exocentric compounds finally stressed (see 3.1 below).

- (13) a. kaba+kulák  
coarse ear  
'mumps'
- b. kıl kuyrúk (may also have regular stress)  
hair tail  
'shabby person'

### 2.3.2 SUB exocentric finally stressed compounds

These are termed *başıbozuk* constructions by Lewis (1967). The lexical item *başıbozuk* 'subversive', 'anarchist' (derogatory) is itself a compound where the subordinator is on the left, the literal translation of which is 'its head is corrupt'.

These compounds fall into the same pattern of having stress on the subordinate constituent, but since this constituent is on the right, these can be mistaken as having word stress. Some examples are given below:

- (14) a. göz-ü      pék  
eye-POSS      strong  
'courageous'
- b. baş-ı      boz-úk  
head-POSS      destroy-A  
'subversive', 'anarchist' (derogatory)
- c. el-i      maşa-lí  
hand-POSS      thong-A  
'authoritarian woman'

The structure of these compounds, [[N+POSS]+A] is, interestingly, also the structure of a particular type of AP, illustrated in (15a). Such APs have a corresponding paraphrase, [A+[N+sufA]], illustrated in (15b):

- (15) a. dam-ı kırmızı      b. kırmızı dam-lı  
roof-POSS red      red roof-ADJ  
'red roofed'      'red roofed'

The pair of AP constructions in (15) and the compounds in (14) raise the question as to what the head of these compounds are. Since they are adjectival constructions, their heads should be adjectives, i.e. the overt or covert functional elements that render these adjectival constructions. The only such overt functional form in these is *-lı* in (15b). as such, the examples in (14) and (15a) lack an overt adjecitiviser.

### 2.3.3 Double-headed finally stressed compounds

Most such compounds also contain verbs ((16a) and (16b)), but may also contain other parts of speech (16c):

- (16) a. biç-er            döv-ér  
           reap-PRTC    beat-PRTC  
           ‘thresher’
- b. al-ış            ver-ış  
           buy-VN        sell-VN  
           ‘shopping’
- c. ana            kíz  
           mother        daughter  
           ‘mother-daughter’ (e.g. as in mother-daughter relationship)

### 2.3.4 Finally stressed compounds borrowed from Persian<sup>4</sup>

Persian loans most of which are obsolescent are also finally stressed. These are right-headed where the head is mostly a deverbal adjective:

- (17) a. musiki şínás  
           music know(er)  
           ‘well-versed in music’
- b. vatan pervér  
           country lover  
           ‘patriot’

The two remaining types of finally stressed compound, compounds borrowed from Arabic and native compounds which are left headed are discussed in section 3.2 below.

## 3. Headedness

As mentioned in 1.2, the vast majority of compounds in Turkish are right-headed. However, there are compounds which divert from this generalisation. These are the following types:

- (i) Headless compounds
- (ii) Left-headed compounds

<sup>4</sup> The Arabic and Persian loan words were commonly and productively used in official and literary documents of the Ottoman Empire which was dismantled in the first quarter of the twentieth century. Late Otoman Turkish was also in use in formal and semi-formal speech, certainly in the erudite registers. However, as part and parcel of official and unofficial attempts of forming a national ‘Turkish’ identity, there were efforts to Turkify the lexicon. These movements came in waves, starting from the late nineteenth century, culminating in the 1930s in the Language Reform (see Lewis 2001). As a result, foreign words (by which was meant those of Arabic and Persian origin) were replaced with their Turkish (sometimes revived) equivalents or neologisms created by official committees. The Turkish of today still contains many Arabic and Persian words (as well as French and other Indo-European based ones), but the productive creation of compounds from such sources has ended.

- (iii) Double-headed compounds
- (iv) Variant compounds

### 3.1 Headless compounds

Exocentric compounds may be regularly stressed (18a) or finally stressed (18b) and they may or may not have an LE ((18a) and (18b) respectively):

- (18) a. hanım+el+i  
 lady+hand+LE  
 ‘honey suckle’
- b. gel+gıt  
 come go  
 ‘tide’<sup>5</sup>

### 3.2 Left-headed compounds

Left-headed compounds in Turkish are of two different sources. One of these is the group of Arabic loans which are still in usage:

- (19) a. tebdil-i kıyafet  
 change-LE outfit  
 ‘(dress) in cognito’
- b. arz-ı hal (spelt *arzuhal*)  
 supply-LE condition  
 ‘petition’

(For the LE in these compounds, see section 6.6).

Much more prevalent to an understanding of Turkish compounding are the native left-headed compounds in Turkish. There are structurally of the type [N+A]+LE. Although in the majority of cases the item that the LE attaches to is a noun, in these compounds LE attaches to an adjective, as in (20a). However, it is not clear that the element bearing the LE is the head in these constructions, as such compounds can be paraphrased as [A+N] compounds as in (20b):

- |      |   |   |
|------|---|---|
| (20) | a. dut kuru-su<br>mullberry dry-LE<br>‘dried mullberries’ | b. kuru dut<br>dry mullberry<br>‘dried mullberries’ |
|------|---|---|

One might think that the adjective *kuru* ‘dry’ is used as a noun (20a) (see section 6). However the word *kuru* does not exist independently as a noun. Nor is it possible to paraphrase all as [A+N] compounds (as illustrated in (20b)) in this way:

- |      |                                      |                                 |
|------|--------------------------------------|---------------------------------|
| (21) | a. çingene pembe-si<br>Gypsy pink-LE | b. *pembe çingene<br>pink Gypsy |
|------|--------------------------------------|---------------------------------|

<sup>5</sup> See Dede 1978: 7-8 for different types of exocentric compound.



‘bright pink’

In fact, the structures in (20a) and (20b) normally yield two different compounds:

- (22) a. elma yeşil-i                                  b. yeşil elma  
        apple green-LE                              green apple  
        ‘apple green’                                ‘green apple’

This brings to question the headedness of (20a). Either the construction is left-headed, or the adjective (e.g. *kuru* ‘dry’) is reanalysed as a noun. We leave this point for further research.

### 3.3. Double-headed compounds

As discussed in 2.3.2, double-headed compounds are usually finally stressed, and are usually nouns that can be used as adverbs:

- (23) gelin kaynana  
        daughter-in-law mother-in-law  
        ‘(as) daughter-in-law (and) mother-in-law’

### 3.4 Variant compounds

There are a group of compounds in Turkish where the order of the constituents is variant:

- (24) a. balık ızgara                                  b. ızgara balık  
        fish grill                                      grill fish  
        ‘grilled fish’                                ‘grilled fish’

We call such compounds *variant compounds*. All of these refer to dishes, and then to a small subgroup of them. In these compounds, the position of the head is not clear, nor is the relation between the constituents. Moreover both constituents in either variant can be stressed. The difficulty in understanding the structure of (24) is compounded by the fact that (i) many compounds indicating dishes are not variant and have a fixed order of constituents as in (25), and (ii) this fixed order is not uniform, as in (26):

- (25) a. biber dol-ma                                  \*dolma biber  
        pepper stuff-N  
        ‘stuffed capsicums’  
        b. patates tava                                  \*tava patates  
        potato pan  
        ‘pan(-fried) potatoes’
- (26) a. kuzu fırın                                      \* fırın kuzu  
        lamb oven                                      oven lamb  
        ‘oven (-baked) lamb’  
        b. \*sütlaç fırın                                  fırın sütlaç  
        milk-pudding oven                          oven milk-pudding  
        ‘oven (-baked) milk-pudding’

## 4. Compounding by reduplication: a derivational operation

Full reduplication of a stem functions as a category changing operation in Turkish. We consider these constructions compounds on the grounds that they are formed of two stems and they have the typical stress of compounds (cf. repetition compounds in Fabb 1998):

- (27) yavaş yavaş  
slow slow  
'slowly'

### 4.1 The reduplication of nouns and adjectives

Reduplication of adjectives and nouns forms adverbs:

- (28) a. zaman zaman (doubling of N)  
time time  
'now and again'
- b. sabah sabah (doubling of N)  
morning morning  
'in the morning'. 'at the crack of dawn'
- c. güzel güzel (doubling of A)  
nice nice  
'nicely'
- d. tane tane (doubling of classifier)  
item item  
'separately'

Reduplication may also create individuating adjectives:

- (29) kırmızı kırmızı  
red red  
'(individuated instantiations of) red'

### 4.2 The reduplication of bound stems

In Turkish, not only free stems but also bound stems may be reduplicated. There are two types of bound stem which undergo such doubling.

#### 4.2.1 Onomateopoeic words

A somewhat productive type of compounding involves a group of bound stems, some of which are onomateopoeic words (i.e. nouns, as they name sounds). Some examples are *şıkır-* (the sounds of beads or jewellery), *şırıl-* (the sound of water), *zırl-* (the sound of sobbing),

*takır-* (the sound of knocking on wood), *bangır-* (the sound of a thump). These stems either require a derivational suffix, e.g. the verb forming suffix *-dA* (hence, e.g. *takır-da* ‘clatter’), or they require full reduplication in order to surface (see also Bauer 2003). These form mostly manner adverbs but also adjectives:

- (30) a. zırıl zırıl (ADV)  
‘in a manner of sobbing’  
  
b. takır takır (both ADV and A)  
‘in a clatter’, ‘clattering’

#### 4.2.2. Other bound stems

There is a group of bound stems which, when fully reduplicated, also form manner adverbs, but unlike the ones in 4.2.1. compounding is the only means in which these bound stems surface:

- (31) haldır haldır  
‘in the manner of ...’ check dict.!

A related type are pairs of bound stems:

- (32) a. paldır küldür  
‘in an unprepared fashion’  
  
b. dangıl dungul  
‘in a coarse manner’

## 5. The N/A distinction and its implications for the SUB/ATT distinction

As is widely acknowledged in the literature, operational means for unambiguously distinguishing nouns from adjectives in Turkish are difficult to find. While many adjectives can bear nominal inflectional forms, many nouns can be used for naming properties, hence are gradable like adjectives (Braun and Haig 2000, Uygun 2007). This is one of the prime factors that makes it difficult to ascribe a structural description to compounds along the SUB/ATT divide as this description hinges upon whether a particular item can uniquely function as an attribute or not.

### 5.1 The blurred border between Nouns and Adjectives

Like in other languages, nouns in Turkish can be used attributively (see Dede 1978, Gökdayı 2007):

- (33) a. demir kapı  
iron door  
‘iron door’  
  
b. ipek çorap  
silk stocking  
‘silk stocking’

c. altın yüzük  
gold ring  
'golden ring'

Since the non-head of such constructions denote the material that the item in the head position is made up of, these [N+N] compounds are ATT. In other words, an ATT compound does not necessarily have an adjective as its non-head.

The blurred border between nouns and adjectives in Turkish presents problems with respect to the divide concerning the SUB/ATT distinction, as nominal inflection can be attached to adjectives:

(34) kırmızı-lar-dan  
red-PL-ABL  
'of the red ones'

Further, some adjectives can be used for denoting the item that they describe, e.g. *genç* is 'young' but also '(a) young person'.

Conversely, a noun denoting an entity can be used for naming the typical property of that entity, e.g. *kadın* is 'woman' but also 'female'. Hence such 'nouns' sometimes appear in the comparative and superlative forms:

(35) a. kadın doktor  
woman doctor  
'female doctor'

b. ...-dan daha kadın  
-ABL than woman  
'more feminine than...'

(adapted from Braun and Haig 2000)

Moreover, even the special compound suffix (see section 6), can occur on an adjective that cannot be used as a noun in other contexts:<sup>6</sup>

(36) genç iri-si  
young stout-LE  
'plump teenager'

However, it is not possible to deduce that nouns and adjectives are therefore indistinguishable and that all forms are ambiguous between a noun and an adjective:

- (i) Predicative adjectives and some complex adjectives do not denote entities, as some nouns cannot refer to properties (see Uygun 2007).
- (ii) Adjectives used as nouns are lexicalised in meaning, i.e. when the adjective 'güzel' is used as a noun it has a lexicalised and animate reference: 'beautiful woman (or

---

<sup>6</sup> *İri* 'stout' is a predicative adjective which cannot be used as a noun:

\*bir iri  
a stout

man)', and not e.g. beautiful house, book, etc.. . Hence such terms cannot refer to just any object which has the property described by the adjective.

- (iii) Although adjectives carry nominal inflectional morphemes as in (34), they are not 'inflected' with nominal inflection. As there is no overt pronominal form denoting Ns in Turkish (e.g. 'one' in English) and as NPs can be headless, the inflectional markers appearing morphologically on a noun can appear on an adjective where there is no N head in the construction. Thus adjectives appear to be inflected for number, person and case, making them look like nouns.
- (iv) There are also some suffixes which mark a structure either as an adjective or a noun. We now turn to these and see how far these take us in understanding whether a particular constituent is a N or A and in covering some ground towards the SUB/ATT divide.

## 5.2 Overt affixes as partial cues

Turkish has derivational suffixes which mark denominal adjectives (e.g. *-li*, *-sal<sup>7</sup>*). and deadjectival nouns (e.g. *-lik*). The presence of such suffixes gives some clue as to the syntactic class of an item, and thus provides a basis on which the SUB/ATT distinction can be made. The presence of the LE *-si* (see section 6) is also one factor that may be used for a similar distinction, as *-si* cannot be used where the dependent constituent is a property.

- (i) the denominal adjective markers *-sal* and *-li<sup>8</sup>*:

When *-sal* appears on a non-head, then the compound is ATT:

- (37) bitki-sel hayat (ATT)  
plant-A life  
'vegetative state', 'coma'

Similarly, *-li<sup>9</sup>* can be used for forming compounds as well:

- (38) söz-lü tarih (ATT)  
word-A history  
'oral history'

What is noteworthy is that these suffixes are not productive as parts of compounds:

- (39) \*bitki-sel çay (cf. bitki çay-ı, see () below)  
plant-A tea  
(Int.int.: herbal tea)

<sup>7</sup> See Göksel 1990 for the formal properties of *-sal*.

<sup>8</sup> Neither of these suffixes unambiguously form compounds, of course, since they are adjectival suffixes, they can be used productively for modifying nouns inside NPs:

- (i) bitki-li çay  
plant-A tea  
'tea with herbs in it'

<sup>9</sup> The suffix *-li* also forms nouns: *atlı* 'horseman', *nişanlı* 'fiancee'.



man barber  
'male barber'

man barber-LE  
'barber for men'

It is also interesting that in *-si* compounds where the non-head is an adjective, such adjectives are not used attributively:

(42) a. *tirşe reng-i*  
turquoise colour-LE  
'the colour turquoise'

b. *'iki' sayısı*  
two numeral-LE  
'the numeral 'two''

Here rather than defining a property of the head (i.e. a subset within the denotation of the head (e.g. *kırmızı X*: the (set or instantiation of a single) *X* such that it has the property 'red'), it names the category it belongs to. The adjective here is used as a name given to a property. The same is true of the examples below:

(43) a. *kırmızı kelime-si*  
red word-LE  
'the word 'red''

b. *A harfi*  
A letter-LE  
'the letter 'A''

c. *gül çiçeğ-i*  
rose flower-si  
'the flower rose'

d. *kız çocuğ-u*  
girl child  
'girl'

An extension of this function of *-si* is prevalent in constructions where the non-head denotes the 'name' given to the head (cf. Dede 1978:130-133):

(44) *Atatürk Bulvar-ı*  
Atatürk Boulevard-LE  
'Atatürk Boulevard'

However, problems regarding the category of the non-head in constructions with *-si* arise in two cases:

- (i) when the type of the item denoted by the head and the non-head are indistinguishable (6.2)
- (ii) when *-si* is optional (6.3)

## 6.2 *-si* as a diagnostic of syntactic categories?

In some compounds with *-si*, the non-head both deliniates the type of item denoted by the head, but is also is part of the make-up of the item that the compound refers to, a typical property of attributive compounds:

- (45) bitki çay-ı                      (SUB)  
      plant tea-LE  
      ‘herbal tea’

While ‘herbal tea’ is a tea made of herbs and is not itself a herb (a typical property of SUB compounds), it is also a substance which would not exist without the contribution of the non-head (a typical property of ATT compounds). The same applies to other such compounds, see Bisetto and Scalise 2006).

## 6.3. The optionality of *-si*

It was mentioned above that the presence or lack of *-si* was crucial in identifying SUB and ATT compounds (see 30 above). However, there are certain cases where *-si* is optional. These are the following:

- (i) Terms for dishes

- (46) a. patlıcan dol-ma-sı                      b. patlıcan dolma  
      eggplant stuff-N-LE                      eggplant stuff-n  
      ‘stuffed eggplants’                      ‘stuffed eggplants’

- (47) a. kabak kalye-si                      b. kabak kalye  
      courgette stew-LE                      courgette stew  
      ‘courgette stew’                      ‘courgette stew’

- (ii) Street names (composed only with *sokak* ‘street’)

- (48) a. Köşk Sokağ-ı                      b. Köşk Sokak  
      Köşk street-LE                      Köşk street  
      ‘Köşk Street’                      ‘Köşk Street’

- (iii) Some terms for colours

- (49) bayrak kırmızı-sı                      kan kırmızı  
      flag red-LE                      blood red  
      ‘cadmium red’                      ‘crimson’

It should be noted that what looks like the optionality of *-si* may sometimes yield two different constructions, especially when the first item is a term relating to nationality. Here constructions without *-si* are APs:



*Compound*

*AP*

- (50) a. Alman general-i  
German general-LE  
'German general'
- b. Alman general  
German general  
'German general'

This is due to the ambiguous terms denoting nationalities and ethnicities, more explicit in the example below:

- (51) a. Türk asker-i  
Turkish soldier-LE  
'The soldier of the Turkish army'
- b. Türk asker  
Turkish soldier  
'a soldier who is Turkish'

See section 8 for further discussion.

#### **6.4 The place of *-si* inside a word:**

There are two properties of *-si* which are interesting from a morphological point of view. One of these is that it can occur only once on a stem, even in those cases where it is syntactically and semantically required. The second one is that it is a closing suffix. We take each one in turn.

##### **6.4.1 *-si* can only occur once on a stem**

There are potentially two cases where *-si* should be able to occur on a stem more than once :

- (i) Where both instantiations of *-si* are compound markers:

This would be an embedded compound construction:

- (52) \* [ Z [ X Y *-si*] *-si*]

Instead, the following structure is grammatical:

- (53) Polonya gölge tiyatrosu(\*-su) (cf. *Polonya tiyatrosu* and *Gölge tiyatrosu*)  
Poland shadow theatre-LE  
'Polish shadow theatre'

These are discussed in section 8 in more detail under recursivity..

- (ii) Where one marks the compound and the second one marks possession:

One of the functions of *-si* is as a third person possessive marker. Where both a compound marker and the possessive *-si* are semantically required, only one (the possessive marker) surfaces (Lewis 1962, Dede 1978, Kornfilt 1984, Göksel 1988, 1993, Schroeder 1999, Van Schaik 2001):

- (54) a. diş fırça-sı                      b. fırça-m                      c. \*diş fırça-sı-m                      d. diş fırça-m  
           tooth brush-LE                      brush-1<sub>POSS</sub>                      tooth brush-LE-1<sub>POSS</sub>                      tooth brush-1<sub>POSS</sub>  
           ‘tooth brush’                      ‘my brush’                      (Int: my tooth brush)                      ‘my tooth brush’

#### 6.4.2 -si is a closing suffix

-*si* has to occur last within the functional elements of a word, excluding, of course the functional elements that link the word to higher structures, e.g. case (see Ralli 2007)). This property of -*si* has been noted to form bracketing paradoxes (Göksel 1989, 1993), illustrated in (55) and (56). The plural form of -*si* compounds are expected to display *si+ler* sequences, whereas the grammatical form is *ler+si* (= *leri*):

- (55) a. balo elbise-si                      b. \*balo elbise-si-ler                      c. balo elbise-ler-i  
           ball dress-LE                      ball dress-LE-PL                      ball dress-PL-LE  
           ‘ball gown’                      (Int: ball gowns)                      ‘ball gowns’

-*si* can also not precede derivational suffixes even though the compositionally ‘correct’ position might be as such, hence *lık+si* (= *lığı*), and *ci+si*:

- (56) a. balo elbise-si                      b. elbise-ci                      c. \*balo elbise-si-ci                      d. balo elbise-ci-si  
           ball dress-LE                      dress-AG                      ball dress-LE-AG                      ball dress-AG-LE  
           ‘ball gown’                      ‘dress shop/vendor’                      ‘vendor for ball gowns’

#### 6.5 The linking element -li

-*li* (which we refer to as -*li* for reasons of simplicity) is a productive denominal adjective suffix:

- (57) bahçe-li  
           garden-A  
           ‘with a garden’

Some N+*li* combinations can only occur as parts (heads) of compounds, hence it functions as a linker of two constituents:

- (58) a. sarı saç-li                      b. \*? saç-li  
           yellow hair-a  
           ‘blond haired’
- a. kısa etek-li                      b. \*? etek-li  
           short hemline  
           ‘short hemmed’

#### 6.6 The Arabic linking element in borrowed compounds

This element is part of the obsolescent Arabic compounds and is rarely used productively in Modern Turkish:

- (59) ehl-i keyf  
           folk-LE pleasure

‘hedonistic’

## 7. Higher order compounds

Higher order compounds are productive compounds which have sentence type elements as non-heads and higher order nouns as heads:<sup>12</sup>

- (60) a. [biz adam olmayız] düşünce-si  
[we will never improve] idea-LE  
‘the idea that we will never improve’
- b. [Cumhurbaşkanımız kim olacak] soru-su  
[who will be our president] question-LE  
‘the question as to who our president will be’
- c. [tanrı var mıdır] soru-su  
[does God exist] question-LE  
‘the question ‘does God exist’
- d. [ay ne kötü] çığlığ-ı  
[how awful] shriek-LE  
‘the shriek ‘how awful’

The construction is obligatorily inflected by the LE *-si*, indicating that these are *-si* compounds par excellence, with the only difference being that their heads are higher order predicates, usually deverbal nouns, but also other types<sup>13</sup> (see Schaaik 2002 for a detailed analysis of such compounds).

Higher order nouns can take finite complements as in (60). These may be declaratives (60a), wh-questions (60b), yes/no questions (60c) and exclamations (60d). They can also take non-finite complements, inflected with one of the markers of non-finiteness (see Göksel & Kerslake 2005), underlined in the example below:

- (61) [Söylenenlerin büyük bir yalan olduğu] iddia-sı  
that what is said is a big lie claim-LE  
‘The claim that what is said is a big lie’

Overall, higher order compounds are:

- (i) marked with *-si* on the head
- (ii) right-headed
- (iii) regular with respect to stress
- (iv) have higher order nouns as heads
- (v) have phrasal constituents, sometimes full Ss as non-heads

<sup>12</sup> These are similar to the post-lexical compounds discussed in Shibatani and Kageyama (1988). Here we adopt the terminology of van Schaaik.

<sup>13</sup> As Schaaik (2002) discusses, nominals that show ‘things’ or ‘persons’ refer to ‘first order entities’. Higher order compounds are based on a nominal head of some higher order entity expressed by various types of propositional verbs such as those that denote intellectual or emotional attitude, mental perception and other propositional terms, e.g. *believe, fear, feel, claim, thought, wish, hope, question, answer*, etc.



paper that those compounds analyzed as right-branching ones do not appear to be instances of recursivity on the right, as they do not seem to be compatible with expected characteristics of compounds. (68) presents an example analyzed as a RBC in (Spencer 1991 Schaaik 2002, Özsoy 2004).

- (68) Türk Tarih Kurum-u  
 Turkish history institution  
 Institute of Turkish history

For Spencer (1991) and Schaaik (2002), (68) is an example of a right-branching compound in which *tarih* ‘history’ plus *kurum* ‘institution’ form the compound *tarih kurum*. This predicate is combined with *Türk* ‘Turkish’ as well as the compound marker *-si* (surfacing here as *-u*). According to this analysis, the main difference between right and left-branching compounds is that the former do not carry the compound marker *-si* and start at a lower level with a term and a nominal which is itself a compound on its own,<sup>14</sup> as in the case of the above example *tarih kurum-u* ‘history institution’.

A closer examination of (68), however, reveals that *Türk tarih kurumu* has two underlying structures, yielding different interpretations.<sup>15</sup>

- (69) a. [Türk [tarih kurum]-u]                      b. [[Türk tarih] kurum]-u  
 Turkish history institution                      Turkish history association  
 ‘History Association of Turkey’                      ‘Association of Turkish history’

While (69a) refers to an institution where all types of history may be investigated, (69b) refers to an institution which studies the history of Turkey or Turks. Note also that (69b) seems to be a truncated form of :

- (70) [[Türk tarih]-i kurum]-u  
 ‘Association of Turkish history’

on a par with (64), displaying left branching.

One diagnostic test which might give us a clue as to whether (68) is also right branching concerns the stress pattern of the above compound. As has been previously discussed (Section 2.1), compounds in Turkish are stressed on the stressable syllable of the non-head. In a compound such as *tarih kurumu*, the stress is on the non-head *tarih* ‘history’. If the analysis that *Türk tarih kurumu* is right-branching were on the right track, one would expect to find stress to appear on *Türk*. However, stress in (68) is on *tarih*. If *Türk* were the non-head of the larger compound, it would have taken stress. As such, we claim that in (69a) *Türk* must be an adjective. Note that in (70) *Türk* is stressed, and it is the non-head of the lower compound. The question remains however, why the ambiguity in (68) is not resolved by stress, if stress is such a strong indicator of structure. As such, we cannot explain why the head (within the non-head) is stressed in (69b) rather than the lower non-head.

## 8.2 Variant adjectival modification

<sup>14</sup> See König (1987), Hankamer (1988) and Göksel (1989) for opposing views .

<sup>15</sup> Spencer (1991) suggests that those as in (69b) are not a possible structures in view of the presence of structures such as (70). Especially with terms for institutions, this seems to be an option, though.

In Turkish compounds, adjectives modifying the compound are placed to the left of the compound, as in (71):

- (71) a. yaşlı ceviz ağac-ı            \*ceviz yaşlı ağac-ı  
      old walnut tree-LE  
      ‘old walnut tree’
- b. küçük sokak çocuğ-u        \*sokak küçük çocuğ-u  
      small street child-LE  
      ‘young street urchin’

Adjectives, thus, do not interfere between the two constituents of the compound, providing evidence for its inseparable nature. Similarly other constituents (e.g. clitics) cannot occur inside a compound.

It should be noted, however, that the last two decades has seen a change in this pattern and the occurrence of an adjective inside a compound for the modification of the head is possible to a limited extent (e.g. see Dede 1978:30 where the unacceptability of this construction is discussed. See also Hayasi 1996, Özsoy 2004, among others for the structure of this type of adjectival modification.)

- (72) a. Adalet Bakan-ı  
      justice minister-LE  
      ‘Minister of Justice’
- b. *eski* adalet bakan-ı  
      old justice minister-LE  
      ‘former Minister of Justice’
- c. adalet *eski* bakan-ı  
      justice old minister-LE  
      ‘former Minister of Justice’

As pointed out in Özsoy (2004), this particular use is limited and is not productive.

## 9. Inflection

Inflectional elements marking the compound as a whole occur on the rightmost constituent (but inside the LE *-si*, see section 6.2):

- (73) a. ipek çorap-lar  
      silk stocking-PL  
      ‘silk stockings’
- b. kadın çorap-lar-ı  
      woman stocking-PL-LE  
      ‘women’s stockings’

### 9.1 Inflected constituents of compounds

Both constituents of compounds can be complex, a natural result of the agglutinative nature of the language. The complexity may be in the form of an inflected verb (74a-b), an inflected noun (74c-f) or a participle (74e,f) (see Dede 1978:59):

- (74) a. kap-tı kaç-tı (snatch-P run.away-P) ‘public car’  
 b. çit kır-ıl-dı-m (sW break-PASS-P-1SG) ‘mummy’s girl/boy’  
 c. göz-ü pek (eye-3POSS-1SG sturdy) ‘courageous’  
 d. kar-dan adam (snow-ABL man) ‘snowman’  
 e. gün-e bak-an (sun-DAT look-PRTC) ‘sunflower’  
 f. kör-ler okul-u (blind-PL school-LE) ‘school for the blind’  
 f. yan-ar dağ (burn-PRTC mountain) ‘volcano’

## References

- Bauer, L. (2003). (3<sup>rd</sup> imprint) *Introducing Linguistic Morphology*. Edinburgh: Edinburgh University Press.
- Beard, R. (1998). ‘Derivation’, in A. Spencer and A.M. Zwicky (eds.) *Handbook of Morphology*, 44-65. Oxford: Blackwell.
- Bisetto, A. and S. Scalise (2005). The classification of compounds. *Lingue e Linguaggio*, 4/2.
- Braun, F. & Haig, G. (2000). ‘The noun/adjective distinction in Turkish: An empirical approach’, in A. Göksel and C. Kerslake (eds.) *Studies on Turkish and Turkic Languages*. Wiesbaden: Harrassowitz.
- Charette, M. , A. Göksel and S. Şener (2007). Initial stress in morphologically complex words in Turkish: the interface of prosodic structure and ‘phrase’ structure. Paper presented at the Workshop on Phonological Domains, Universals and Deviations, DGfS-29, Siegen University. To be submitted to J. Grijzenhout and B. Kabak (eds.) *Phonological Domains; Universals and Deviations*. Interface Explorations, Mouton de Gruyter.
- Dede, M. (1978). A Syntactic and Semantic Analysis of Turkish Nominal Compounds. PhD dissertation, University of Michigan.
- Fabb, N. (1998). ‘Compounding’, in A. Spencer and A.M. Zwicky (eds.) *Handbook of Morphology*, 66-83. Oxford: Blackwell.
- Gökdayı, H. (2007). ‘Takısız Ad Tamlamaları’ in *Turcology in Turkey; Selected Papers*, Laszlo Karoly (ed.), Papers selected by Nurettin Demir and Emine Yılmaz, *Studia Uralo-Altaica* 47, 243-252. Szeged.
- Göksel (1988). Bracketing paradoxes in Turkish nominal compounds. In S. Koç (ed.) *Studies on Turkish Linguistics*, 287-298. Ankara: METU Press.
- Göksel, A. (1990). Semantic properties of Turkish nominal compounds. *Proceedings of the Tenth Annual Conference on Greek Linguistics*, 379-398. Thessaloniki: University of Thessaloniki Press.
- Göksel, A. (1993). Levels of representation and argument structure in Turkish. PhD dissertation, SOAS, University of London.
- Göksel, A. & C. Kerslake. (2005). *Turkish. A Comprehensive Grammar*. London: Routledge.
- Hankamer, J. (1988). Parsing nominal compounds in Turkish. Morphology as a computational problem. *Occasional Paper 7*. UCLA.
- İskender, H.İ. (2007). Vowel deletion in Turkish. Boğaziçi university ms.
- Hayasi, T. (1996). ‘The dual status of possessive compounds in modern Turkish’, in Á. Berta, B. Brendemoen, & C. Schönig (eds.) *Symbolae Turcologicae*, Vol. 6, 119-129. Uppsala.

- Kabak, B. (2007). Hiatus resolution in Turkish: an underspecification account. *Lingua* 117/8:1378-1411.
- Kabak, B. & A. Revithiadou (2006). The phonology of clitic groups: Evidence from Turkish and Asia Minor Greek. Paper presented at the Thirteenth International conference on Turkish Linguistics. Uppsala university.
- Kabak, B. and A. Revithiadou (2007). Prosodic structure above the phonological word. Paper presented at the Workshop on Phonological Domains; Universals and Deviations, DGfS-29, Siegen University.
- Kabak, B. & Vogel, I. (2001). The phonological word and stress assignment in Turkish. *Phonology* 18, 315-360.
- Kornfilt, J. (1984). 'The 'Stuttering Prohibition' and morpheme deletion in Turkish' in E. Erguvanlı-Taylan and A. Aksu-Koç (eds.) Proceedings of the Second international Conference on Turkish Linguistics. İstanbul: Boğaziçi University Publications.
- König, W.D (1987). 'Nominal compounds in Turkish'. *Zeitschrift für Sprachwissenschaft* 6: 2. 165-185.
- Lees, R. (1961). *The phonology of modern standard Turkish*. Uralic and Altaic Series, V. 6. Bloomington: Indiana University Publications.
- Lewis, G.L. (1967). *Turkish Grammar*. Oxford: Clarendon Press.
- Lewis, G.L. (2001). *The Turkish Language Reform. A Catastrophic Success*. Oxford, Oxford University Press.
- Özsoy, S. (2004). 'Dışışleri eski bakanı ve Türkçe'nin yeni yapısı', in Z. Toksa (ed.) *Kaf Dağı'nın ötesine varmak, Günay Kut Armağanı*, Vol. 3, *Journal of Turkish Studies*, Türklük Bilgisi Araştırmaları 28/1. Harvard University.
- Ralli, A. (2007). Compound markers and parametric variation. To be published in *Typology*.
- van Schaik, G. (2002). *The Noun in Turkish. Its Argument Structure and the Compounding Straitjacket*. Wiesbaden: Harrassowitz Verlag.
- Schroeder, C. (1999). *The Turkish Nominal Phrase in Spoken Discourse*. Wiesbaden: Harrassowitz Verlag.
- Shibatani, M. and T. Kageyama (1988). 'Word formation in a modular theory of grammar: a case of post-syntactic compounds in Japanese'. *Language* 64, 451-84.
- Spencer, A. (1991). *Morphological theory*. Oxford: Blackwell.
- Uygun, D. (2007) Lexical Categories in Turkish. Boğaziçi University ms.