

Giorgio Francesco Arcodia* and Isa Sari

Towards a typology of coordinating compounds in Turkish

<https://doi.org/10.1515/flin-2018-0008>

Submitted November 07, 2016; Revision invited August 13, 2017;

Revision received September 6, 2017; Accepted November 23, 2017

Abstract: Although there is a rich literature on compounding in Turkish, coordinating compounds have not received enough attention from researchers in the field of morphology. The Turkish lexicon is characterised by a large number of coordinating compounds, belonging to different subclasses; this paper is a first attempt to sketch a typology of these word forms. We will first provide an overview of the different types of coordinating compounds in the language, exploring the limits of this phenomenon, and we will then analyse them in terms of five main parameters: (a) the semantic relation between the constituents and the referent of the whole compound; (b) input and output word classes; (c) morphological cohesiveness; (d) prosodic cohesiveness; (e) degree of conventionalisation/lexicalisation. We will argue that there is a fundamental divide between endocentric and exocentric coordinating compounds; the latter appear as more lexicalised and entrenched in the Turkish lexicon than the former and also seem to be closer to prototypical words in the morphology-syntax continuum.

Keywords: compounding, coordinating compound, Turkish, word-formation

1 Introduction

Compounding is one of the most productive mechanisms for naming concepts and enriching the vocabulary in the world's languages; it can be described simply as the combination of two or more lexical items to form a new one. A compound, that is, the output of the mechanism, may be defined as a "lexical unit made up of two or more elements, each of which can function as a lexeme independent of the other(s) in other contexts" (Bauer 2001: 695). This simple definition, however,

*Corresponding author: **Giorgio Francesco Arcodia**, University of Milano-Bicocca, Department of Human Sciences for Education, Piazza dell'Ateneo Nuovo, 1, 20126, Milan, Italy, E-mail: giorgio.arcodia@unimib.it

Isa Sari, Hitit University, Faculty of Arts and Sciences, Department of Turkish Language and Literature, İkbal Kent Kampusu, 19040, Merkez/Çorum, Turkey, E-mail: isasari@hitit.edu.tr

raises two immediate problems: first, the problem of the input of compounding and, second, the definition of compounds as opposed to other “multi-word expressions” (Bauer 2001: 704). Bauer (2006: 719) opts for the term ‘subword’ as an all-encompassing term for the possible basic units of compounding: “(...) the forms in which the individual subwords appear may be differently defined in different languages; a citation form in one, a stem in another, a specific compounding form in yet a third, a word form in a fourth.” Thus, the delimitation of the category of compounds, as opposed to derived words and phrases, somehow depends on the morphological profile of the language at issue. Moreover, not everybody agrees on the morphological nature of compounding (see Scalise and Vogel 2010). According to Bağrıaçık and Ralli (2015), compounds may be morphologically or syntactically built; the *locus* in which compounds are generated depends on the language at issue, and even on the specific subtype of compound. As we will see below, Turkish coordinating compounding may be located on different points along the morphology-syntax continuum, depending on the pattern.

In Turkish, the basic units of compounding can be roughly defined as words, most often in their bare (citation) form; however, compound constituents may also contain derivational and inflectional suffixes, and compounds built on bound stems are also attested (Göksel 2009). Turkish is rich in derivational morphemes (Kornfilt 1997: 445) and, while word formation via suffixation is frequent and typologically distinctive for the language, compounding has also been used productively as a word formation strategy since the first attestations of Old Turkic (e.g. *kışı oğl-ı* ‘human child-CM, mankind’, *adak ağng* ‘foot soreness’, *ög kañg* ‘mother father, parents’; Poppe 1966; Erdal 2004: 381–383; Sertkaya 2006: 110; Tekin 2006: 24; Eraslan 2012: 137).¹

Compounds have been classified in different ways, based on several criteria, such as presence or absence of a structural/semantic head or the grammatical relation between the constituents (see Bisetto and Scalise 2005; Scalise and Bisetto 2009; for extensive overviews). Within this line of research, coordinating compounds (also known as ‘copulative’, Bloomfield 1933; ‘co-compounds’, Wälchli 2005; ‘coordinate compounds’ Scalise and Bisetto 2009) have had a special position due to the symmetric relation between the constituents since the times of the early Sanskrit grammarians, who coined the time-honoured label

¹ The glosses follow the general guidelines of the *Leipzig Glossing Rules*; additional glosses include AOR ‘aorist’, CM ‘compound marker’ and NEC ‘necessitative mood’. Unless otherwise specified, all examples are from Turkish. Since hyphenation conventions vary for some Turkish compounds, whenever both hyphenated and non-hyphenated forms are used we will put the hyphen between brackets (‘(-)’).

dvandva (Ten Hacken 2000: 358) – a term which has since then been applied to Western languages (cf. i.a. Marchand 1969). Nevertheless, when compared to other types of compounding, coordinating compounds have often received less attention in the literature (Wälchli 2005), even though recent research has shown that coordination in compounding is productive also in European languages, including English (see Olsen 2001: 279, 312 for a discussion). In Turkish as well, coordinating compounding is productive to a degree, and there are a great deal of attested examples of this type of complex word forms, but despite the rich literature on compounding in Turkish (e.g. König 1987; Kornfilt 1997; Van Schaaik 2002; Göksel and Haznedar 2007; Göksel 2009), coordinating compounds have so far not received much attention, with the exception of some typological studies (see Wälchli 2005; for some related examples from Turkish, and compare Göksel 2009) and a recent paper on Verb-Verb compounds (Güneş 2009).

In this paper, we will provide an overview of coordinating compounds in Turkish, classifying them according to five main parameters: (a) the semantic relation between the constituents and the referent of the whole compound; (b) input and output word classes; (c) morphological cohesiveness; (d) prosodic cohesiveness; (e) degree of conventionalisation/lexicalisation. We will argue that there is a rather strong divide between two (macro-)classes of coordinating compounds, namely endocentric compounds, such as *oyuncu-yönetmen* ‘actor-director’, and exocentric compounds, such as *ana(-)baba* ‘mother-father, parents’: while the former appear to be semantically transparent and have consistent properties (in terms of semantics, *locus* of inflection, stress pattern, and word classes), the latter form a less coherent class, with significant differences among subclasses and even individual items, and appear to be more lexicalised. Although it has been claimed that endocentric coordinating compounds are restricted to the so-called Standard Average European (SAE) linguistic area (Arcodia et al. 2010), a fair amount of examples are attested in Turkish, and the pattern appears to be productive. While exocentric coordinating compounds have always been part of the Turkic lexicon (Wälchli 2005: 198–199), endocentric compounds could be a more recent development, possibly following the model of SAE languages.

The paper is organised as follows. First, we discuss the definition of the term ‘coordinating compound’ used in the present study, introducing the discussion on the relation between coordinating compounds and other complex word forms in the Turkish lexicon (Section 2). Next, we will sketch an analysis of coordinating compounds, based on the five parameters introduced above (Section 3). Finally, we will summarise the main points of the paper and give some hints for further research (Section 4).

2 Defining coordinating compounds

The definition of coordinating compounds involves two aspects: the identification of compounds as opposed to phrases and other constructions, and the definition of coordination as opposed to asymmetrical relations. In this section, we will introduce these two issues and we will keep referring to them when discussing our parameters.

2.1 Compounds, phrases and related constructions

As already mentioned in the introduction, the delimitation of compounding is a notoriously thorny issue in the study of linguistic morphology, and Turkish is no exception to this. Take, for instance, what is arguably the most productive pattern of compounding in Turkish (Göksel 2009), termed either ‘possessive compounds’ or simply ‘compounds’ (Bağrıaçık and Ralli 2015): these are asymmetric (attributive or subordinate) right-headed compounds which contain the compound marker $-(s)I(n)$,² as in e.g. *elma ağac-ı* ‘apple tree’. These compounds have several properties in common with phrases, as e.g. in constituent coordination (Göksel 2009; Bağrıaçık and Ralli 2015), and are considered by some to be syntactically generated (Bağrıaçık and Ralli 2015; Kamali and İkizoğlu 2015); however, they also have several properties of words (e.g. non-reversibility of the constituents), and hence, some authors (e.g. Kunduracı and Göksel 2016) maintain that they are morphological objects.

Needless to say, a thorough discussion on the issue of the borderline between compounding and syntax in Turkish is well beyond the scope of the present article (the reader may refer to Göksel 2009; Bağrıaçık and Ralli 2015; Kunduracı and Göksel 2016; *inter alios*). However, for the purposes of our discussion, we need to (coarsely) distinguish at least coordinating compounds from NP coordination. We will lay the basis for this distinction here, but, as said above, we will be getting back to this in the rest of the article.

Given the difficulties and the uncertainties in this domain, which led to the lack of consensus as sketched above, in this paper, we chose to adopt a ‘continuum’ view of the morphology-syntax boundary, in which there may be items that have more ‘word’ features and items that have more phrasal features. Starting

² Göksel and Haznedar (2007) treat $-(s)I(n)$ as a linking element; some believe that it should be analysed as ‘3SG.POSS’ (e.g. Johanson 1998: 50; Ergin 2011: 382–383; Banguoğlu 2011: 334). However, in modern Turkish linguistics, it is mostly simply termed ‘compound marker’ (Kornfilt 1997: 474).

from Bauer's (2001) broad definition of 'compounding' quoted in the introduction, we will discuss the 'compoundness' of the constructions at issue here, based on the well-known Lexical Integrity Hypothesis (cf. Bağrıaçık and Ralli 2015; Kunduracı and Göksel 2016; for a recent overview, see Lieber and Scalise 2006). Specifically, we focus on a few common diagnostics for 'wordhood', namely: word atomicity (the impossibility of having contextual inflectional exponents, functional and/or lexical morphemes within the compound), prosodic cohesion (having a single primary stress), stress pattern (lexical vs. phrasal stress), and fixed order of constituents; to this we add recursive expandability and non-compositionality of meaning, which are diagnostics for lexicalisation (and hence, in a sense, lexemic status). This will be done in Sections 3.3–3.5.

2.2 Coordination in compounding

Coordinating compounds may be loosely defined as complex word forms in which the constituent lexemes are in a coordination relation (see Renner 2008).³ In coordinating compounds, no subordination or attribution relation is observed between the constituents,⁴ and all the constituents have equal weight semantically (Lieber 2005: 378; Arcodia et al. 2010: 177). While some scholars suggest that coordinating compounds do not have an overt semantic head (Booij 2005: 80; Ten Hacken 2000: 359), many others believe that the constituents of a coordinating compound do show head-like functions (Fabb 1998: 66–67; Scalise and Guevara 2006: 191; Kageyama 2010), as the meaning of the whole compound may be considered to be, in a sense, the 'sum' of its constituents. In other words, two (or more) semantic heads can be posited since all the constituents equally contribute to the meaning of the whole (see Haspelmath 2002: 89; Scalise and Bisetto 2009: 46). However, Bisetto and Scalise (2005) convincingly argue that coordinating compounds may actually be endocentric or exocentric. Let us examine again the examples mentioned in the Introduction:

- (1) *oyuncu-yönetmen*
actor-director
'actor-director'

³ On the definition of coordination, see e.g. Haspelmath (2004: 33–37).

⁴ Here we follow Bisetto and Scalise's (2005) classification of compounds, according to which attributive compounds are those in which a constituent indicates a property or attribute of the other, as in Eng. *blue cheese* or *snail mail*, and subordinate compounds are those in which there is a complement relation, as in Eng. *taxi driver* or *pickpocket*.

- (2) *ana(-)baba*
 mother-father
 ‘parents’

In (1), the compound has two heads: an ‘actor-director’ is both an ‘actor’ (*oyuncu*) and a ‘director’ (*yönetmen*); in other words, it is endocentric. The resulting compound is in fact a hyponym of its constituents (see Arcodia et al. 2010). In (2), however, the relationship between the referent of the whole compound and the constituents is of a different nature: ‘parents’ is not something that is both a ‘mother’ (*ana*) and a ‘father’ (*baba*). ‘Parents’ is a superordinate concept with respect to ‘mother’ and ‘father’ (as e.g. in *ana-baba hakkı* ‘the right of parents’). The compound is therefore exocentric. Both subtypes are attested in Turkish, as (1)–(2) exemplify, but as we shall see below, they have different characteristics.

Needless to say, the very notion of headedness in word formation is controversial (for an overview, see Scalise and Fábregas 2010; Arcodia 2012): for compounding, much weight has been given to the ‘hyponymy’ criterion (i.e. a compound is a hyponym of its head), to the ‘categorical’ criterion (i.e. the head is the category determinant; Andreou 2014), and also to the *locus* of inflection/derivation (which should be the head). A thorough discussion of this *vexata quaestio* is obviously beyond the scope of the present paper; hence, we shall limit ourselves to a few basic remarks on headedness in Turkish coordinating compounds. The first two criteria seem to apply to endocentric compounds: in *oyuncu-yönetmen*, the compound is a hyponym both of *oyuncu* ‘actor’ and of *yönetmen* ‘director’, and both constituents are also category determinants; however, as we shall see in greater detail below, overt morphological markers generally appear on the right-hand constituent only. We believe that this marking pattern is expected if compounds are word forms and, hence, show lexical integrity (the same generally holds e.g. for English: *actor-directors*; see below, Section 3.3). For exocentric coordinating compounds, the hyponymy criterion fails to apply, as expected: *ana(-)baba* is neither a hyponym of *ana* nor of *baba*; here, both constituents may be claimed to be category determinants, but this is not the case for other exocentric coordinating compounds, such as e.g. *gel-git* ‘come-go, tide’, in which two verb stems form a noun. In point of fact, as we shall see below (Section 3.2), most exocentric coordinating compounds in Turkish are nouns, but they are often formed by verbal or even adjectival constituents, and their categorial output is sometimes unpredictable. Moreover, as is apparent from an example like *gel-git* ‘come-go, tide’, the meaning of the compound may not be easily recovered from the meanings of its constituents (see below, Sections 3.1, 3.5). Hence, we believe it is safe to say that they are exocentric both from a categorial and from a semantic point of view.

While the distinction between exocentric coordinating compounds and asymmetric (attributive/subordinate) compounds is usually quite straightforward, for endocentric coordinating compounds this distinction may be tricky. For instance, Eng. *woman doctor* is listed in Bisetto and Scalise (2005) as an example of a coordinating compound; however, the relationship between the constituents does not appear to be perfectly symmetrical, but rather one of modification/attribution (a doctor who belongs to the female sex). Thus, the inversion of its constituents (*doctor woman*) arguably does not produce a word with the same meaning. As to Turkish, endocentric coordinating compounds may also have an attributive reading: an *oyuncu-yönetmen* can be both an ‘actor-director’ and ‘a director, who however has something of an actor’. In fact, given the fact that the categories ‘noun’ and ‘adjective’ are not clearly distinguished in Turkish (see e.g. Braun and Haig 2000; Van Schaaik 2002), and given also that these constructions have the same stress pattern as non-coordinating right-headed compounds (see below, Section 3.4), they have been analysed as right-headed attributive compounds (see Van Schaaik 2002; Göksel 2009). Braun and Haig (2000) propose that the word class of ‘nominals’ in Turkish is a continuum between prototypical nouns and prototypical adjectives, with many items located on different points of the continuum. On the basis of an experiment with acceptability judgements, they show that speakers consider as equally acceptable both order of constituents for noun-noun compounds in which both constituents are at the same level of ‘nouniness’ (e.g. *kadın polis* ‘woman police officer’) but prefer the order with the ‘more adjectival’ element on the left if there is a difference between the two (e.g. *genç polis* ‘young police officer’).

In a similar spirit, we believe that ‘prototypical’ endocentric coordinating compounds are made of constituents that are not only on the same level of ‘nouniness’ (a somewhat vague notion) but also on the same taxonomic level: thus, in our view, *oyuncu-yönetmen* ‘actor-director’ is a better candidate than *kadın polis* ‘woman police officer’, as ‘actor’ and ‘director’ are on the same taxonomic level, whereas ‘woman’ is a more general term than ‘police’ and is not in the same semantic area (see Arcodia forthcoming; we will get back to this in Section 3.1).

3 A classification of Turkish coordinating compounds

As stated in the introduction, in this section we will propose a classification of Turkish coordinating compounds based on five main parameters, comparing the characteristics of exocentric and endocentric compounds. As mentioned above,

endocentric compounds emerge as a homogeneous and consistent class, distinct from that of exocentric compounds, and are less lexicalised; the latter, on the other hand, are not as consistent, and appear to be more lexicalised, and closer to the prototypical word.

3.1 The semantic relation between the constituents and the referent of the whole compound

Generally speaking, coordinating compounds can be a combination of synonyms, antonyms or related notions (Fabb 1998; Wälchli 2005). As to exocentric compounds, in Turkish they may consist of constituents which share the same meaning ('synonymic co-compounds' in Wälchli 2005 terminology), as e.g. in *şan şöret* 'fame fame, glory'. Coordinating compounds built on constituents with a synonymous or near-synonymous relation between the constituents often denote (roughly) the same concept as the constituents themselves; however, the semantics of the whole construction may differ from its constituents and show superordinate features to a degree, as in e.g. *ar namus* 'shyness rectitude, purity' or *düğün dernek* 'wedding organisation, festival'. Coordinating compounds made of antonyms are also possible in Turkish, like *yap-boz* 'construct-ruin, jigsaw', *gel-git* 'come-go, tide'; they may indicate generalising semantics, i.e. showing both the meanings of their constituents, as in the above mentioned *yapboz*. Lastly, constituents with closely related meanings (parallel notions) may together form coordinating compounds, mostly belonging to the 'collective' type (Wälchli 2005) like *eş dost* 'partner fellow, acquaintances', *para pul* 'money money, assets', or *yer gök* 'land sky, world' (see below); collective compounds denote a set comprising all referents having the properties shared by the compound constituents (Wälchli 2005: 141).

Some exocentric coordinating compounds may have been formed with lexemes which are not used as free ones in Standard Turkish (akin to what Wälchli 2005 termed 'ornamental compounds') as in *kap kakak* 'pot ?, utensils', *takım taklavat* 'equipment ?, paraphernalia' or *börtü böcek* '? insect, creepy-crawlies'. Such coordinating compounds mostly denote superordinate notions and plurality, encompassing more than two related entities. Coordinating compounds with constituents bearing no clear semantics are also attested, like *abur cubur* '? ?, junk food'.

As mentioned before, it is not possible to observe a subordinate relation in coordinating compounds (Renner 2008: 608): there is no semantic dominance of one constituent over the other(s); every constituent of a coordinating compound contributes to the meaning of the whole compound. Cross-linguistically, it has

been observed that the semantic relationship between the constituents can be either of ‘and-coordination’ (combination) or of ‘or-coordination’ (alternative) (see Bauer 2011; Olsen 2015); in Turkish, however, virtually only ‘and-coordination’ is attested in compounding, both exocentric and endocentric.⁵ As to exocentric compounds, when verbal constituents are involved, one may get two different subtypes of combination: (i) concatenations where two actions are carried out simultaneously, e.g. *uyu-r-gez-er* ‘sleep-AOR.PTCP-ramble-AOR.PTCP, sleepwalker’, and (ii) concatenations which involve successive states of affairs, e.g. *al-iş-ver-iş* ‘get-NMLZ-give-NMLZ, shopping’, which requires both ‘giving (money)’ and ‘getting (a product)’. Note that in both cases, the meaning of the whole is based on the sum of the meanings of the constituents, and no contrast or alternative seems to be implied (compare e.g. Mandarin Chinese *sheng-fū* ‘win-lose, victory or defeat’).

Among coordinating constructions, a somewhat particular case is that of constructions like the following:

- (3) *ana(-)kız*
 mother-daughter
 ‘mother and daughter’ (Göksel 2009: 221)

The form *ana(-)kız* ‘mother and daughter’ looks pretty much identical to *ana(-)baba* ‘parents’ discussed above. Göksel (2009) states that pairs such as *ana(-)kız* are indeed compounds because they follow the final stress pattern of Turkish words, as opposed to the left-hand stress pattern of phrases (*infra*, Section 3.4). However, *ana(-)kız* is significantly different from *ana(-)baba*: whereas the latter actually means ‘parents’ and can be used as a generic term, *ana(-)kız* always indicates a specific pair of mother and daughter, not a superordinate category (in its nominal use; see [8] below). In other words, *ana(-)kız* is always used referentially; compare (4), (5), (6) and (7):

- (4) *ana-kız* *kitap oku-du.*
 mother-daughter book read-PST
 ‘the mother and her daughter read a book (together)’
- (5) *ana-baba* *ol-mak kolay değil.*
 mother-father be-INF easy NEG.COP
 ‘it’s not easy to be parents’

⁵ Although collocations like *uzun kısa* ‘long or short’ and *büyük küçük* ‘large or small’ are attested (we thank an anonymous reviewer for pointing this out and for providing examples).

- (6) [?]*ana-kız* *ol-mak kolay değil.*
 mother-daughter be-INF easy NEG.COP
 ‘(intended meaning:) it’s not easy to be a mother and a daughter’
- (7) **ana-baba kitap oku-du.*
 mother-father book read-PST
 ‘the mother and father read a book (together)’

In (4), *ana(-)kız* is used referentially, to indicate a specific pair of mother and daughter. In (5), *ana-baba* generically refers to ‘parents’. On the other hand, *ana (-)kız* is at best odd if used non-referentially, as in (6); in turn, *ana-baba* is plainly ungrammatical if used referentially as in (7).⁶ However, an apparently non-referential use of *ana(-)kız* is actually possible when this construction is used as an adverbial modifier rather than as a noun⁷:

- (8) *biz ana kız gel-eceğ-iz.*
 we mother daughter come-FUT-1PL
 ‘We will come alone, as mother and daughter’

Hence, the restriction to referential usages applies to *ana(-)kız* when used as a noun, the prototypical referential item; we may remark here that, generally speaking, when nouns are used as modifiers they tend to shed their referential properties.

The restrictions described above for *ana(-)kız* also apply to other similar items (e.g. *gelin kaynana* ‘daughter-in-law and mother-in law’; Göksel and Haznedar 2007). *Ana(-)kız* is the instantiation of a fairly regular pattern, in which the kinship term denoting the ‘older’ member in the relationship comes first and the other kinship term denoting the ‘younger’ comes last: other

⁶ Note, however, that when used in its plural form, *ana-baba* may also be referential. Compare:

- (i) *ana_baba-lar çocuk-ları ile tartış-ma-malı.*
 parents-PL child-3PL.POSS with argue-NEG-NEC
 ‘Parents shouldn’t argue with their children.’
- (ii) *ana_baba-lar okul-da buluş-tu.*
 parents-PL school-LOC meet-PST
 ‘Parents met at the school.’

While in (i) *ana babalar* is non-referential, in (ii) it can have specific reference. This, however, is never possible for the singular *ana-baba*.

⁷ We would like to thank an anonymous reviewer for pointing out this example to us.

examples include *baba(-)oğul* ‘father-son’, *amca(-)yeğen* ‘uncle-nephew/niece’, *abi(-)kardeş* ‘elder brother-younger sibling’, *abla(-)kardeş* ‘elder sister-younger sibling’. A further requirement for this construction is that the relationship between the members be ‘direct’ (for lack of a better term): there has to be an ‘of’ relation between the constituents (i.e. X is the X of Y, and Y is the Y of X).⁸ Hence, possible pairs such as **ana-yeğen* ‘mother-nephew/niece’ or **amca-ana* ‘uncle-mother’ are ruled out. From a constructionist perspective, this pattern may be regarded as a constructional schema, with two variable slots specified for the above mentioned semantic features; the individual items created by means of this construction are somehow ‘between’ morphology and syntax, depending on their nominal or adverbial nature (usage?), but their relative degree of fixity and the productivity of the schema means that they qualify as ‘multi-word word-like expressions’ (Booij 2010; compare Masini 2006; treatment of binomials); in their adverbial usage, they qualify as an exocentric [N N]ADV compounding pattern (see the next section).

For endocentric coordinating compounds, the relation between the constituents is a straightforward coordination of related items; some compounds made of antonyms are attested as well, such as *acı tatlı* ‘bitter sweet, bittersweet’. As indicated above (Section 2.2), we believe that when the left-hand constituent of the compound is not on the same taxonomic level and not in the same semantic area as the second one, the compound leans towards an attributive interpretation. This is not to say that a clear borderline may be drawn but, rather, that the plausibility of a coordinating interpretation is higher for items like *oyuncu-yönetmen* ‘actor-director’, the coordination of two labels for related professional roles, or *sağır-dilsiz* ‘deaf-mute’, the coordination of two (often cooccurring) disabilities; on the other hand, a coordinating interpretation is less plausible e.g. for *kadın polis* ‘woman police officer’, or for *çocuk gelin* ‘child bride’. The latter two examples are, in our view, ambiguous between a coordinating and an attributive reading.

As to the relation between the constituents and the whole compound, we stated above (Section 2.2) that the basic difference between endocentric and exocentric coordinating compounds lies in the relationship between the constituents and the compound. An exocentric coordinating compound like *çekyat* ‘sofa bed’ consists of two lexical items: *çek* ‘to pull’ and *yat* ‘to lie down’, but the output meaning is lexicalised to denote a tool. Both the meanings of *çek* and *yat* can be found in the semantics (or functions) of *çekyat*, but the result is rather exocentric since the output is a noun, and the meaning of the compound is not fully predictable from the meaning of its parts (although this is not

⁸ We would like to thank an anonymous reviewer for pointing this out to us.

necessarily always the case; see below, Section 3.5). Endocentric compounds, on the other hand, are fully compositional, and no reference to an external head is needed for their interpretation: an *ambulans uçak* ‘ambulance aircraft’ is both an ‘ambulance’ and an ‘aircraft’, and no extra semantic element is involved. In endocentric coordinating compounds, each constituent independently denotes different attributes or aspects of the whole construction. Marchand defines this type of compound as ‘appositional’ (1969: 61–62), while Olsen uses the label ‘coordinative appositional’, stating that the subtype of coordination involved is ‘accidental’ coordination (as opposed to ‘natural’ coordination; Olsen 2015: 368). Generally speaking, the main restriction on the input for this class of compounds is that the lexemes to be combined are understood as referring to a property of an entity, rather than to the entity *per se*: thus, for instance, pairs like the above mentioned exocentric compound *yer gök* ‘land sky, world’ do not seem to be very suitable for a property reading (something which is both land and sky?). In endocentric coordinate compounds, what is being combined are, in essence, properties. In the case of adjectives like e.g. *acı tatlı* ‘bittersweet’, this is obvious; in the case of nouns, perhaps less so, but we do believe that it is properties that are combined (following Arcodia forthcoming): for instance, an *actor-singer* is someone who can both act and sing professionally. A property reading, however, is not normally available for verbs, which are in fact used to form exocentric compounds, and not endocentric ones.

Finally, despite the rather clear distinction between endocentric and exocentric coordinating compounds, one potentially ambiguous case is that of compounds made of coordinate aorist participle stems (see the next section), such as *biç-er-döv-er* ‘mow-AOR.PTCP-beat-AOR.PTCP, harvester’ or *uyu-r-gez-er* ‘sleep-AOR.PTCP-ramble-AOR.PTCP, sleepwalker’. We list them among exocentric compounds, but (as an anonymous reviewer pointed out) they could be interpreted as endocentric compounds describing ‘a machine that mows and ties’, or ‘a person that sleeps and rambles’, just like an endocentric compound such as *araştırmacı yazar* ‘researcher writer, someone who does both research and writing’. However, note that the deverbal (aorist participle) constituents of these compounds are not generally used on their own (with some exceptions), and hence the word class and other head features should not come from these constituents. Thus, in order to properly interpret them, reference to an implicit head is needed: *makine* ‘machine’ for *biçerdöver* ‘harvester’ (compare the subordinating compound *biçer döver makinesi* ‘harvester machine’) and *kışi* ‘person’ for *uyurgezer* ‘sleepwalker’ (compare Eng. *redhead*). We thus maintain that they belong to the exocentric coordinating class.

3.2 Input and output word classes

Besides the close semantic relationship that usually holds between them, the constituents of Turkish coordinating compounds also show strong structural similarities with each other. Constituents of a Turkish coordinating compound are items belonging to the same word class, and, in the case of exocentric compounds at least, they often have the same affixes; that is, some kind of structural harmony is observed in most coordinating compounds.

For exocentric compounds, the following patterns are attested:

- (9) Noun-Noun
iş güç
 work strength
 ‘routine’
- (10) Verb-Verb⁹
tut-kal
 hold-stand
 ‘glue’
- (11) Verb-PST-Verb-PST
in-di-bin-di
 get.off-PST-get.on-PST
 ‘minimum fare’
- (12) Verb-AOR.PTCP-Verb-AOR.PTCP
biç-er-bağla-r
 mow-AOR.PTCP-tie-AOR.PTCP
 ‘reaper machine’
- (13) Verb-NMLZ-Verb-NMLZ
ye-me-iç-me
 eat-NMLZ-drink-NMLZ
 ‘eating and drinking (food and beverage)’

⁹ Actually, the basic stem of a Turkish verb is identical to the second person singular imperative (Kornfilt 1997: 40–41); however, since there does not seem to be any command implied here, we chose to gloss it as a bare stem.

The first three patterns presented here make use of the bare stem form; as to (11), here the verb is used in a past tense form. This *-DX* past tense suffix is sometimes used as a nominalizer as in *girdi* ‘input’ (from the verb *gir* ‘to enter’), or *çıkıtı* ‘output’ (from the verb *çık* ‘to exit’). The patterns in (12) and (13) also require further clarification. As to Verb-AOR.PTCP-Verb-AOR.PTCP compounds (12), we mentioned in the preceding section that their constituents are nominalised, but they are not generally used on their own (with some exceptions) as nouns. We gloss them as verbs here, but their status is ambiguous between noun and verb; they are listed among compounds with verbal constituents in Göksel and Haznedar (2007). As to (13), the nominalised forms are termed ‘verbal nouns’ in Turkish grammar (in the example, *-mA*; see Göksel and Kerslake 2005: 87) and they generally describe activities or states, as in *gel-me* ‘come-NMLZ, coming’ (*Can’ın erken gelmesi* ‘Can’s coming early’; Göksel and Kerslake 2005: 363); verbal nouns are often lexicalised as ordinary nouns indicating concrete referents (e.g. the result of an action/process), e.g. *kıyma* ‘mince-NMLZ, minced meat’.¹⁰ Again, Göksel and Haznedar (2007) include compounds made of verbal nouns among those with verbal constituents, but we feel that treating them as nouns is more appropriate.

In the patterns exemplified above, the output is invariably a noun, independently of the input category; as mentioned earlier (Section 2), this attests to the fundamental (categorical) exocentricity of these constructions. Interestingly, a combination of two nouns may instead yield an adverb, as in *yaz kış* ‘summer winter, continuously’ (Göksel and Haznedar 2007); we also mentioned above the pattern ‘older relative – younger relative’ yielding adverbs (besides nouns), exemplified by *ana(-)kız* ‘(as) mother and daughter’ (see (8) above). In essence, differently from endocentric compounds (see below), the output category of an exocentric compound is not always predictable, although nominal output is overwhelmingly dominant.

As to endocentric coordinating compounds, only the following two combinations are attested:

- (14) Noun-Noun
Araştırmacı yazar
 researcher writer
 ‘researcher-writer’

¹⁰ Another deverbal nominalising suffix used in compounds is *-(y)Iş* (as in *al-ış-ver-iş* ‘get-NMLZ-give-NMLZ, shopping’), which forms verbal nouns indicating either the manner of an action, or a single instance of an event/action (e.g. *gid-iş* ‘go-NMLZ, going, departure’; Göksel and Kerslake 2005: 370–371).

- (15) Adjective-Adjective
acı tath
 bitter sweet
 ‘bitter sweet’

Thus, whereas verbs feature prominently in exocentric compounds, they seem to be absent in endocentric compounding, both as input and as output; as said above, this can be explained by the fact that it is not easy to obtain a property reading for a verb. In endocentric compounds, categorial identity is perfect: nouns beget nouns, and adjectives beget adjectives; that is to say, these word forms have two categorial heads (i.e. they are not exocentric).

3.3 Morphological cohesiveness

As mentioned above, Turkish compounds show a high degree of lexical integrity as inflectional exponents whose scope extends over the compound as a whole generally appear only on the right-hand constituent (Göksel 2009). This is consistently observed in endocentric coordinating compounds:

- (16) a. *araştırmacı yazar-lar*
 researcher writer-PL
 ‘researcher-writers’
 b. *araştırmacı-lar yazar-lar*
 researcher-PL writer-PL
 (i) *‘researcher-writers’
 (ii) ‘researchers and writers’

The plural form in (16b), with double number marking, is ungrammatical as the plural of *araştırmacı yazar* ‘researcher-writer’; it is however acceptable if it indicates ‘researchers and writers’ as two separate groups (i.e. a set of people who are researchers and another set representing writers). In the latter case, we are obviously dealing with NP coordination and not with a compound. The same contrast emerges when we add the coordinating conjunction *ve* ‘and’:

- (17) *araştırmacı ve yazar-lar toplantı-ya katıl-dı.*
 researcher and writer-PL meeting-DAT participate-PST
 (i) ‘(a/the) researcher and (the) writers participated in the meeting.’
 (ii) ‘(the) researchers and (the) writers participated in the meeting’

Even if the plural marker appears only on the second constituent, the presence of the coordinating conjunction turns the construction into a phrase: in (17), the scope of *-lar* may extend over *yazar* ‘writer’ only (i), or over the whole construction (ii), but in both cases we are dealing with separate entities, not with a set of researchers-writers.¹¹

Note, also, that the insertion of a determiner such as *bir* ‘a/an’ between the constituents is possible for endocentric compounds (on asymmetric compounds, see Göksel 2009); however, this makes a coordinative interpretation unacceptable, turning them instead into right-headed attributive constructions. Compare:

- (18) a. *oyuncu-yönetmen*
 actor-director
 ‘actor-director’
 b. *oyuncu bir yönetmen*
 actor a director
 ‘a director who has something of an actor’

Constructions like (18b) are not coordinating constructions; hence, endocentric coordinating compounds do not allow the insertion of a determiner between the constituents (they may, however, appear in front of the compound: *bir oyuncu-yönetmen* ‘an actor-director’, without violating its integrity). In short, endocentric coordinating compounds seem to show word atomicity.

As a general rule, exocentric compounds, just like endocentric compounds, do not allow internal number and case inflections, but (mostly, symmetrical) TAM (i.e. tense-aspect-mood) markers and nominalisers are allowed¹²; they likewise do not allow the insertion of other lexical morphemes, or of the determiner *bir* ‘a/an’. Compare:

- (19) a. *biç-er-döv-er-ler*
 mow-AOR.PTCP-beat-AOR.PTCP-PL
 b. **biç-er-ler-döv-er-ler*
 mow-AOR.PTCP-PL-beat-AOR.PTCP-PL

¹¹ The syntactic phenomenon by which one grammatical ending scopes over parallel words is termed ‘suspended affixation’ in Turkish linguistics (for an overview, see Kabak 2007).

¹² Actually, in endocentric coordinating compounds one finds constituents such as *yazar* ‘writer’ (see (16)–(17) above), in which the aorist participle suffix *-ar* is analogous to *-er* in *biç-er döv-er* ‘harvester’. However, in endocentric compounds, the aorist participle suffix is part of a lexicalised free word; in contrast, *biçer* and *döver* are not lexicalised and may not be used on their own. They exist only as part of this construction pattern. Hence, we did not consider cases such as *yazar* to show compound-internal TAM marking.

- c. **biç-er-ler-döv-er*
 mow-AOR.PTCP-PL-beat-AOR.PTCP
 ‘harvesters’

Only (19a) is acceptable as a plural form for *biçerdöver* ‘harvester’; both (19b), with overt plural marking on both constituents, and (19c), with plural on the first constituent, are ungrammatical.¹³ In point of fact, the *locus* of plural marking may be used to disambiguate between compounds and NP coordination (phrases): when constructions of the ‘*ana(-)kız*’ type (see above, Section 3.1) are used referentially, they require double plural marking (*ana-lar(-)kız-lar*) (Göksel and Haznedar 2007); in a compound like *ana-baba* ‘parents’, on the other hand, plural marking appears only on the second constituent (*ana(-)baba-lar*), whereas double marking (*ana-lar baba-lar*) is possible only if *ana baba* is understood as NP coordination (i.e. ‘mothers and fathers’, taken as separate entities). This, again, shows that constructions like *ana(-)kız* behave as compounds only in their adverbial use, not in their nominal use.

Also in the case of exocentric compounds, the addition of the conjunction *ve* may be used to distinguish between compounds and NP coordination. Compare:

- (20) a. *ana(-)baba ol-mak zor-dur.*
 mother father be-INF difficult-COP
 (i) ‘being parents is difficult’
 (ii) ‘being a mother and a father is difficult’
 b. *ana ve baba ol-mak zor-dur.*
 mother and father be-INF difficult-COP
 ‘Being a mother and a father is difficult’

While (20a) may be ambiguous between a compound reading (i) and an NP-coordination reading (ii), when the conjunction *ve* is present (20b), only the NP-coordination reading is possible, i.e. the construction refers to a mother and a father considered separately.

Thus, Turkish exocentric coordinating compounds show lexical integrity too because they do not accept elements between their constituents. There are, however, a few exocentric coordinating constructions which look like

¹³ This contrasts, for instance, with Romance, in which compounds may show inflection on the lefthand, righthand or both constituents, depending on the position of the head and on the degree of lexicalisation of the compound: see e.g. Italian *nave traghetto* ‘ferry boat’ > *navi traghetto* ‘ferry boats’; *mezzogiorno* ‘midday, noon’ > *mezzogiorni* ‘middays’; *attore cantante* ‘actor director’ > *attori cantanti* ‘actor-directors’.

compounds, but in which (contextual) inflectional affixes are attached to both constituents, as in:

- (21) *mal-m-ı* *mülk-ün-ü* *topla-ıyp* *kaç-tı*.
 goods-3SG.POSS-ACC property-3SG.POSS-ACC pick.up-CVB flee-PST.3SG
 ‘S/he fled picking up her/his possessions.’
- (22) *ev-i* *bark-ı* *terk ed-ıp* *dağ-a* *yerleş-ti-ler*.
 home-ACC dwelling-ACC abandon-CVB mountain-DAT settle-PST-PL
 ‘They abandoned their home and settled on the mountain.’

The distribution of these ‘anomalous’ compounds is not completely random: double inflectional marking happens mostly with compounds that consist of synonymous or near-synonymous constituents, like those in (21) and (22). We take this as evidence of a lower degree of lexicalisation for these constructions: they are not fully lexicalised and, as such, they are semantically transparent, unlike some exocentric compounds (like the above mentioned *biçerdöver* ‘harvester’ and *gelgit* ‘tide’). They are considered by some to be ‘doublets’ (Tur. *ikileme*) rather than true compounds (Hatiboğlu 1981; Banguoğlu 2011), which is consistent with their lack of lexical integrity. However, in the category of doublets ‘true’ (i.e. showing lexical integrity) exocentric compounds may also be included (see Hatiboğlu [1981: 17, 52] for *alışveriş* ‘shopping’ or Banguoğlu [2011: 309] for *ana baba* ‘parents’), and it is therefore unclear whether the term can be used consistently to label the construction type exemplified above in (21)–(22). Note that both *ev bark* and *mal mülk* are often listed as dictionary entries with a specific usage and meaning (i.e. their meaning is, in a sense, ‘more’ than the meaning of their constituents), hence showing some degree of conventionalisation. Also, as we shall see below (Sections 3.4–3.5), their order is fixed, and they seem to have a different (syntactic) stress pattern from ‘canonical’ exocentric compounds. In other words, they are conventionalised/lexicalised coordinating constructions located between morphology and syntax: they are not prototypical morphological compounds, as shown by their tolerance for internal (contextual) inflection, but they show some degree of conventionalisation and fixity, differently from phrases.

3.4 Prosodic cohesiveness

Coordinating compounds, and Turkish compounds in general, do not seem to form a consistent set from the point of view of prosody. As hinted at above, the

basic stress pattern for Turkish asymmetric (subordinate/attributive) compounds is that in which stress falls on the stressable syllable of the non-head left-hand constituent (Göksel 2009; Güneş 2009; Bağrıaçık and Ralli 2015; ex. from Göksel 2009: 220):

- (23) *elmá ağac-ı*
apple tree-CM
'apple tree'

Also, the second constituent may have secondary (compound) stress on the same syllable as in its free usage (as in *okúľ kitabı* 'school book' Kornfilt 1997: 514).¹⁴ As stated above (Section 2), endocentric coordinating compounds (with two constituents) seem to follow the same left-hand stress pattern of attributive compounds such as (23); secondary stress is also audible on the second constituent¹⁵:

- (24) *arařtırmacı yazâr*
researcher writer
'researcher writer'

The basic left-hand stress pattern of Turkish asymmetric (subordinate/attributive) compounds is the stress pattern of phrases as well, and contrasts with that of words, for which the default position of stress is final; this is taken by some as a further indication of the fact that left-stressed compounds are syntactic in nature (e.g. Kamali and İkozöđlu 2015). Exocentric coordinating compounds are said to conform to the final word stress pattern (Göksel 2009: 220)

- (25) *çek-yât*
pull-lie.down
'sofa bed'

Crucially, the stress pattern for coordinating compounds like (25) is distinct from phrases; that is, if the construction has left-hand (phrasal) stress, it turns into a phrase (hence *çék yat* means 'pull and lie down'; Güneş 2009).

¹⁴ We use the acute accent for primary stress, and the grave accent for secondary stress.

¹⁵ Also, for some instances of compounds which are ambiguous between an attributive and a coordinating reading, only the latter usually allows a short pause between the constituents, as in *oyuncu(-)yönetmen* 'actor-director' or *besteci(-)müzisyen* 'composer musician'. This is another difference between endocentric coordinating compounds and attributive compounds on the prosodic level.

Final (word) stress applies not only to exocentric coordinating compounds, but also to asymmetric compounds in which the second constituent is verbal or deverbal (e.g. *bilgi-say-âr* ‘information-count-AOR.PTCP, computer’; Göksel 2009: 220), as well as to some others which we will not deal with here (see Göksel 2009; Güneş 2009; for a detailed discussion). As stated above (Section 3.1), the final word stress pattern applies also to *ana(-)kız*, which does not behave like a compound in other respects.¹⁶ However, if the deverbal constituent is on the left of the compound and it is not the head, then stress is located on the left-hand deverbal constituent (e.g. *ak-âr-su* ‘flow-AOR.PTCP-water, stream’; Göksel 2009: 220). In Güneş’s (2009) analysis, compounds with a deverbal derivational element on the left-hand constituent have left-hand stress, whereas those with a deverbal constituent to the right have final (i.e. word) stress (see the source for her detailed argumentation).

What happens, then, when *both* constituents are deverbal verbs, as in some of our exocentric compounds (see (10)-(11) above)? Göksel and Haznedar (2007) suggest that exocentric coordinating compounds like *al-iş-ver-iş* ‘get-NMLZ-give-NMLZ, shopping’ and *biç-er-döv-er* ‘mow-AOR.PTCP-beat-AOR.PTCP, harvester’ have final stress, and that they belong to the category of ‘double-headed’ rather than exocentric compounds. Since they do not provide a list of items belonging to this category beyond these examples, or a consistent definition for the class, we could not properly test their suggestion; hence, we just conducted a tentative PRAAT analysis of *alışveriş* and *biçerdöver*.¹⁷ Below are the waveforms, intensity and pitch diagrams for the two words¹⁸:

In Figure 1, the highest pitch and the strongest intensity seem to be visible on the second syllable of the first constituent. Although for *alışveriş* there is a rise in intensity also on the first syllable of the second constituent, there is also a fairly clear decrease in pitch when compared to the first constituent. In Figure 2, on the other hand, we again see a rise in intensity on the second syllable of *biçer*, but the pitch of the nucleus vowel is not higher than that of the following

¹⁶ Also, note that there are some instances of compounds that may alternate between initial and final stress, including a seemingly exocentric coordinating one like *kıl kıyruk* ‘hair tail, shabby person’ (Göksel and Haznedar 2007).

¹⁷ It is important to point out here that the use of PRAAT to detect stress patterns has limited validity. This is because stress cannot be easily measured, and PRAAT diagrams do not show all the indicators of stress; here we analyse variations in pitch and intensity, but e.g. vowel length, which has been often used in the literature as an indicator of stress, is not taken into consideration (we would like to thank an anonymous reviewer for pointing this out to us).

¹⁸ The .wav files for the analysis were downloaded from the database *Sesli Türkçe Sözlük* (‘a voiced Turkish dictionary’) of the *Türk Dil Kurumu* (Turkish Language Association; http://www.tdk.gov.tr/index.php?option=com_seslisozluk; accessed October 5, 2016).

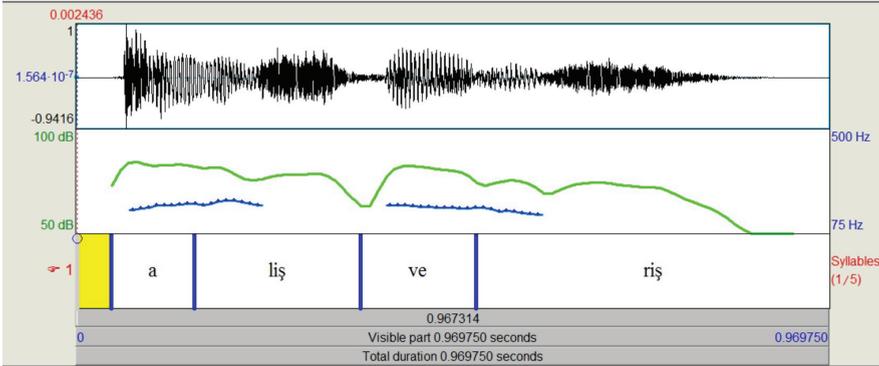


Figure 1: PRAAT analysis for *alışveriş* ‘shopping’.

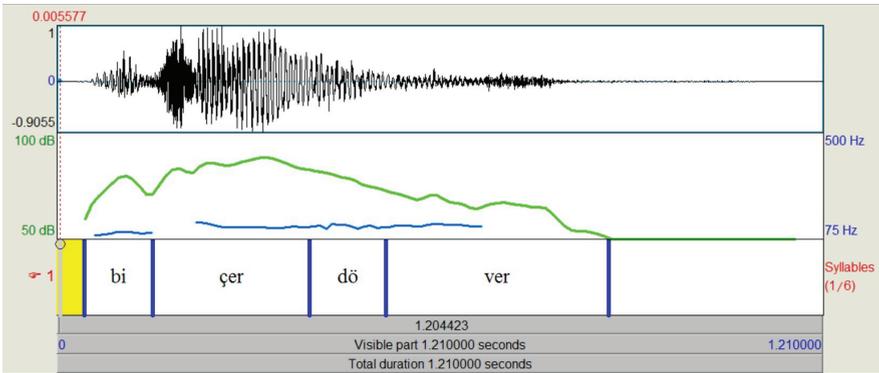


Figure 2: PRAAT analysis for *biçerdöver* ‘harvester’.

vowels. Compare also the PRAAT analysis for *uyu-r-gez-er* ‘sleep-AOR.PTCP-ramble-AOR.PTCP, sleepwalker’, which is analogous to *biçerdöver* as to the number of syllables, inflection and construction of the meaning of the whole compound (i.e. it is an instantiation of the same word formation schema):

When we compare Figures 2 and 3, we may notice that, for *uyurgezer*, the relative prominence in terms of pitch and intensity of the second syllable of the first constituent and the subsequent decrease of the two values are even more obvious.¹⁹ It remains to be seen, therefore, whether the final stress pattern really holds for the class of exocentric and (however defined) ‘double-headed’

¹⁹ Note that, according to Kamali and İkiçoğlu (2015), *uyurgezer* is stressed on the second constituent. Since they provide no pitch diagram, we cannot compare their results to ours.

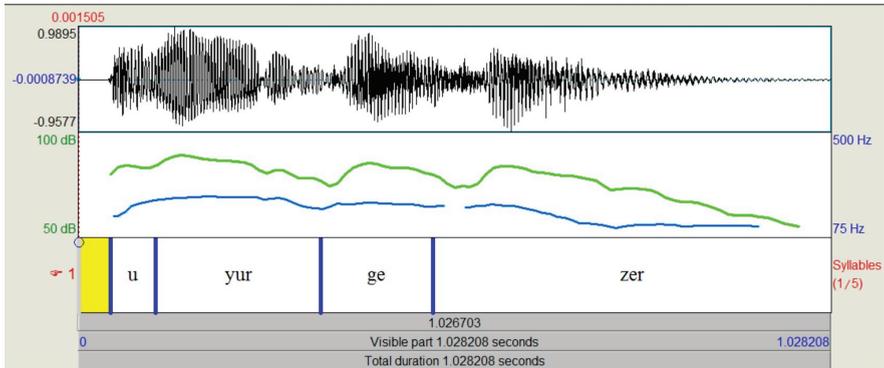


Figure 3: PRAAT analysis for *uyurgezer* ‘sleepwalker’.

compounds: only a thorough survey of a proper sample of compound words could help us answer this question. The tentative conclusion is that coordinating (exocentric) compounds with both deverbal constituents seem to be stressed on the first constituent, differently from other exocentric compounds (i.e. those which do not have a deverbal constituent on the left).

A related issue is the stress pattern of (compound-like) constructions that allow separate number and case marking, like *mal mülk* ‘possessions’ (21) and *ev bark* ‘home’ (22). We argued above (Section 3.3) that they are somehow intermediate between phrasal and word status; if they display a stress pattern unlike that of other compounds, i.e. a less cohesive pattern, it would constitute further evidence for their ‘hybrid’ syntactic/morphological nature. In point of fact, the PRAAT analysis of *mal mülk* (Figure 4) shows no clear difference in pitch and

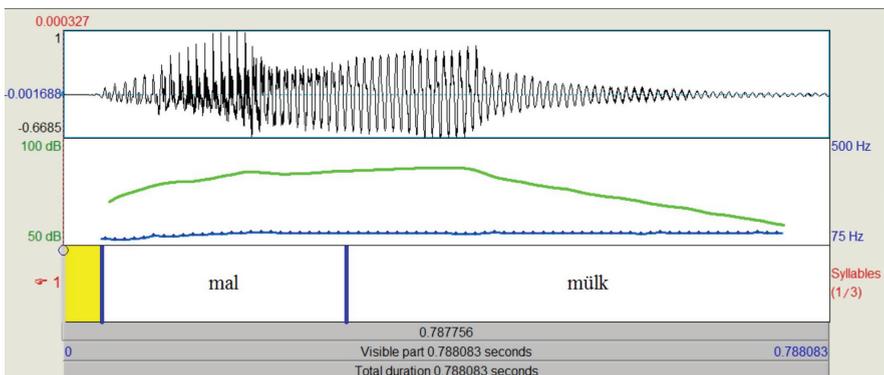


Figure 4: PRAAT analysis for *mal mülk* ‘possessions’.

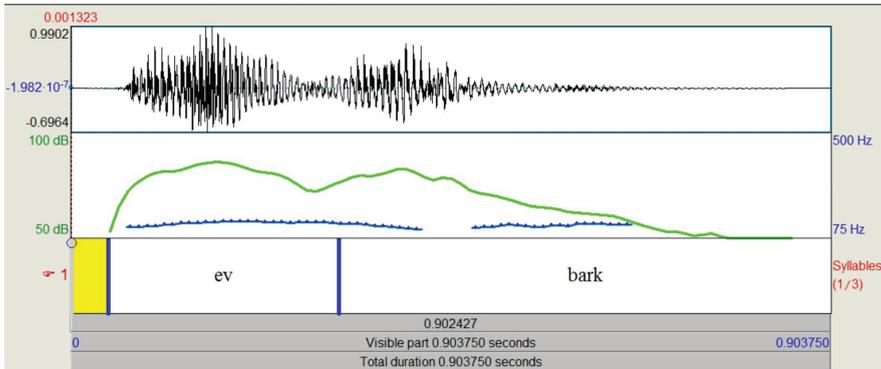


Figure 5: PRAAT analysis for *ev bark* ‘home’.

intensity between the two constituents, as if none of them were ‘stressed’ compared to the other:

The situation is not so clear for *ev bark* ‘home’ (Figure 5), since here the pitch on *ev* appears to be higher than on *bark*; there is also a spike in intensity on the first constituent compared to the second.

Hence, it appears that *mal mülk* ‘possessions’, and perhaps also *ev bark* ‘home’, have a stress pattern which is distinct from that of other exocentric compounds and is more syntax-like. However, note that this is no more than a suggestion for further research; needless to say, a finer-grained analysis of a representative sample of constructions belonging to this subclass is required to make any serious claim.

In short, what emerges from our survey is that endocentric compounds form a seemingly consistent class from the point of view of prosody, and share the same (phrasal) stress patterns as most asymmetric compounds of Turkish. Exocentric compounds, on the other hand, are less coherent: despite the claims in the literature that they consistently have final stress, this might not be true for compounds with both deverbal constituents.

3.5 Degree of conventionalisation/lexicalisation

The degree of conventionalisation/lexicalisation of a compound may be assessed according to two parameters: compositionality of meaning and fixedness of the structure, in terms of stable order of the constituents and recursive expandability. More lexicalised items may be semantically less transparent, have a fixed order of constituents and do not allow expansion by the recursive

addition of other constituents; less lexicalised items, on the other hand, mostly have a fully compositional meaning, may be reversible and may tolerate the addition of other coordinands. Generally speaking, it appears that Turkish exocentric coordinating compounds are more conventionalised/lexicalised than endocentric coordinating compounds.

Firstly, coordinating compounds in Turkish have different degrees of semantic compositionality. As mentioned above, endocentric coordinating compounds are semantically fully transparent: the referent of the whole compound is someone or something which possesses the two identities, properties or characteristics designated by the individual constituent lexemes (be they nouns or adjectives). However, for exocentric coordinating compounds the situation is not so homogeneous.

Exocentric coordinating compounds with synonymous or near-synonymous constituents tend to be the most transparent type, since they normally designate an entity which is somehow the sum of the meanings of its constituent parts; however, some metonymic meaning shift is possible here too, as in *iş güç* ‘work strength, routine’. Metaphoric extensions are also possible, as in *yol yordam* ‘way procedure, good manners’. Also, as is often the case for (exocentric) coordinating compounds, in some cases a ‘semantic integrator’ (Olsen 2014, 2015) has to be posited; in other words, the meaning of the whole compound “is not obtained by simply adding the meanings of the conjuncts together, but rather by subsuming the co-conjuncts under a common conceptual frame” (Olsen 2014: 280). This is the norm for generalising compounds such as *yer gök* ‘land sky, world’, as well as for collective ones, such as *para pul* ‘money money, assets’.

As mentioned above (Section 3.1), with deverbal (and deadjectival) exocentric compounds, it often happens that, in order to understand the compound, reference has to be made to some ‘implicit’ head: compounds designating tools/machines, like *kaz-ar-at-ar* ‘dig-AOR.PTCP-throw-AOR.PTCP, excavator’ seem to follow a fairly regular schema/template, which can be glossed as ‘a machine which does A and B’ (in Booij 2010’s formalism, $[[x]_{V(AOR-PTCP)} [y]_{V(AOR-PTCP)}]_N$ ‘a N which does x and y). In yet other cases, there seems to be a cline from more transparent to less transparent forms although it is actually hard to precisely assess the degree of compositionality of individual forms: compare e.g. *tut-kal* ‘hold-stand, glue’ (10), *in-di-bin-di* ‘get.off-PST-get.on-PST, minimum fare’ (11), *yaz-boz* ‘write-ruin, scratch pad’ and *gel-git* ‘come-go, tide’.

As to the fixedness of the order of constituents, exocentric compounds generally cannot be inverted; this holds not only for ‘canonical’ compounds such as *biçerdöver* ‘harvester’ (**döverbiçer*), but also for ‘hybrid’ forms like *mal mülk* ‘possessions’ (**mül malk*). An exception to this is *kon-ar-göç-er*, ‘settle-AOR.

PTCP-migrate-AOR.PTCP, nomad', also attested as *göçerkonar*; one possible explanation is that two competing version of this compound have lexicalised, which is made possible by its semantics (i.e. the concept of 'nomad' involves 'migrating' and 'settling', but the order does not seem to be crucial).

On the other hand, for most endocentric compounds, like the above mentioned *araştırmacı yazar* 'researcher writer' or *ambulans uçak* 'ambulance aircraft', the order of the constituents may be inverted, as indicated above (Section 2.2). This, again, points towards a lower degree of lexicalisation for endocentric compounds when compared to exocentric compounds. However, note that in other instances a certain order of constituents appears to have been conventionalised also for endocentric compounds as, for instance, in *sağır-dilsiz* 'deaf-mute' and *prens piskopos* 'prince-bishop'. In these two cases, a possible explanation for their apparent fixity is that they are calques from a Standard Average European language with the original order being lexicalised as such. Interestingly, reversibility of the constituents is sometimes possible also for compounds which designate a highly entrenched and conventionalised/lexicalised notion such as *avcı toplayıcı* 'hunter gatherer', attested too as *toplayıcı avcı* 'gatherer hunter'.²⁰

As to recursive expandability, this is generally not allowed for exocentric compounds (e.g. *para pul* 'money money, assets' > **altın para pul* 'gold money money'). On the other hand, endocentric compounds, with the possible exception of the more entrenched/lexicalised (like the above mentioned *sağır-dilsiz* 'deaf-mute'),²¹ may be expanded with the addition of another constituent: *eğitimci araştırmacı yazar şair* 'educator researcher writer poet', for instance, is a possible construction.

3.6 Summary

In the preceding sections, we highlighted the most salient features of different subclasses of coordinating compounds in Turkish. In Table 1, we summarise the main points of our classification, based on the fundamental opposition between endocentric and exocentric compounds.

All in all, what emerges from this survey is that endocentric compounds tend to be, on the whole, more regular and 'predictable' (for lack of a better word) in terms of semantics, *locus* of inflection, stress pattern, word classes,

²⁰ <https://tr.wikipedia.org/wiki/Bu%C5%9Fmanlar> (accessed October 3, 2016).

²¹ An anonymous reviewer pointed out that an expanded *kör sağır dilsiz* 'blind-deaf-mute' is apparently possible; even very conventionalised items, thus, might allow (perhaps limited) recursive expansion.

Table 1: Summary of our classification.

Compound class	Semantic relation between constituents and compound	Input and output word classes	Morphological cohesiveness	Prosodic cohesiveness	Degree oflexicalisation
Exocentric	Additive, collective, generalising, synonymic; reference to a semantic integrator or external head is necessary	All major word classes; different verbal bases; output mostly a noun, sometimes an adverb	Internal number/case inflection or other functional morphology not allowed, but TAM and nominalisers allowed; however, internal number/case inflection for some compound-like coordinating constructions (doublets)	Stress claimed to be on the rightmost constituent, but this may not always be the case (when there is a deverbal constituent on the left); doublets, but not 'true' compounds, might have equal stress on both constituents	Varying degrees of semantic transparency; (near-)synonymic compounds tend to be the most transparent; order of constituents almost never reversible.
Endocentric	Mostly additive coordination of names of professions, places, etc.; each constituent independently denotes attributes or aspects of the whole	Only nouns and adjectives as input; output has the same word class as input	Internal number/case inflection or other functional morphology never allowed	Primary stress on the leftmost constituent as in the case of asymmetric compounds; short pause between the constituents allowed for some items at least	Fully compositional; many items are reversible and recursively expandable

while exocentric compounds represent a complex and varied subclass, with differences among subgroups (and sometimes even for individual items: see the case of *kon-ar-göç-er*, ‘settle-AOR.PTCP-migrate-AOR.PTCP, nomad’ vs. *göçerkonar*).

As to the ‘compoundhood’ (here, wordhood) of each construction type, we may remark that both exocentric compounds and endocentric compounds do not allow internal contextual inflection, although TAM markers are part of some exocentric compounding patterns. As to prosody, exocentric compounds are closer to the prototype of the word since many (but perhaps not all of them) follow the final stress pattern; the opposite happens for endocentric compounds, which follow the stress pattern of phrases (and attributive/subordinate compounds); also, endocentric compounds may sometimes allow a short pause between the constituents. Lastly, while exocentric compounds generally have a fixed order of constituents, are not expandable and may be more or less opaque from a semantic point of view, endocentric compounds are often reversible, may (mostly) be expanded with other coordinands and are fully transparent. Thus, we may say that, on the whole, exocentric compounds are closer to the prototypical word than endocentric compounds: the former score higher in measures of lexical integrity, they are prosodically closer to the word (perhaps with the exception of an initially stressed subclass, see above, Section 3.4) and appear to be more lexicalised (as proven by their fixed structure and semantics). Compound-like constructions like the above mentioned *ev bark* and *mal mülk* (21–22), on the other hand, are further removed from the prototype of the compound(/word): they do show some degree of conventionalisation and fixity but they tolerate internal contextual inflection, just like phrases.

4 Concluding remarks

In this paper, we have shown that Turkish coordinating compounds are a complex and varied class of compounds; what really ties them together is the symmetric nature of the relationship between the constituents. The most basic divide is obviously that between exocentric and endocentric compounds. Generally speaking, endocentric compounds appear to be closer to appositive compounds in several aspects: they always seem to follow the left-hand stress pattern, number and case morphology consistently appears only on the second constituent, and they involve, in a sense, a relation of reciprocal modification (see Olsen 2014: 281). Also, endocentric compounds tend to be more ‘regular’ and homogeneous than exocentric compounds: their meaning is fully

compositional, they are made of constituents belonging to the same word class, which is inherited by the whole compound, and they can be formed only of nouns and adjectives. The only semantic requirement on their constituents is that they denote something that may somehow be interpreted as a property. Lastly, many items in this class allow different orders for their constituents, and recursive expansion, pointing towards a low degree of lexicalisation/conventionalisation; the pattern is productive, which entails both that new items are being created, and that there are several items of relatively recent coinage.

Exocentric compounds, on the other hand, are a more complex subclass, with several patterns, some of which are more regular/idiomatic (e.g. the Verb-AOR.PTCP-Verb-AOR.PTCP pattern, seen in [12]), whereas many items have more idiosyncratic and opaque meanings (like e.g. *gel-git* ‘come-go, tide’). They consistently show word atomicity, with no insertion of functional or lexical morphemes allowed, and no contextual inflection; however, (non-contextual) TAM markers may be present, and they are indeed part of some construction patterns. Their stress pattern may not be entirely consistent, with some items having final stress and others seemingly having initial stress. All in all, what emerges is that exocentric compounds seem to be more entrenched in the Turkish lexicon, and, as such, they exhibit a variety of properties, sometimes inconsistent among subclasses (and even for items in the same subclass), whereas endocentric compounds seem to be less entrenched, more transparent and less lexicalised; also, the endocentric compounding pattern is productive, whereas exocentric compounds do not appear to have significant productivity in the present stage of the Turkish language.

This is consistent with the hypothesis sketched above, according to which endocentric compounds are a relatively recent development in Turkish word formation, probably acquired by calquing from Western European languages.²² Note that a similar situation holds for the Uralic and Baltic languages of Europe, which have fully lexicalised exocentric compounds, often opaque, and less entrenched endocentric coordinating compounds, mostly recent and semantically transparent, probably coined after the model of Western European languages (see e.g. Larsson 2002; Wälchli 2005; Forgács 2015;

²² Note that, since Ancient Greek did have productive endocentric coordinate compounding (e.g. *iatrómantis* ‘healer-diviner’, *hippalektryôn* ‘horse-rooster’; Manolessou and Tsolakidis 2009: 28), a Greek origin of the pattern, predating the contact with Modern European languages, cannot be ruled out in the absence of historical data. However, we believe that this is unlikely, especially since contact with Turkish was more intense at the stage of Medieval Greek (Metin Bağrıaçık, p.c., October 2016), when the exocentric pattern was introduced; moreover, one should not forget the influence of French, a language with productive endocentric coordinate compounding, on Modern Turkish.

Navickaitė-Klišauskienė 2016). Further research on the history of compounding, on the historical and contemporary productivity of the various patterns, and on the stress patterns of different subclasses of compounds may shed further light on the issues raised here.

Lastly, concerning the *locus* of compounding, endocentric compounds have phrasal (left-hand) stress, are fully compositional, often allow for the inversion of constituents and show some recursion, but they do not allow overt morphological elements between their constituents. Exocentric compounds are, however, closer to the prototype of the word, as they have mostly lexical (final) stress, may have non-compositional meanings, (almost) never allow inversion of the constituents and recursive expansion, and show atomicity. Thus, we may say that, in the continuum between syntax and morphology, endocentric compounds are somewhat closer to the former than exocentric compounds; nevertheless, they may be clearly distinguished from NP coordination, and hence are surely not ‘true’ syntactic structures.

References

- Andreou, Marios. 2014. *Headedness in word formation and lexical semantics: Evidence from Italiot and Cypriot*. Patras: University of Patras dissertation.
- Arcodia, Giorgio Francesco. 2012. Constructions and headedness in derivation and compounding. *Morphology* 22. 365–397.
- Arcodia, Giorgio Francesco. forthcoming. Coordinating nominal compounds: Universal vs. areal tendencies. *Linguistics* 56(6).
- Arcodia, Giorgio Francesco, Nicola Grandi & Bernhard Wälchli. 2010. Coordination in compounding. In Sergio Scalise & Irene Vogel (eds.), *Cross-disciplinary issues in compounding*, 177–198. Amsterdam: John Benjamins.
- Bağrıaçık, Metin & Angela Ralli. 2015. Phrasal vs. morphological compounds: Insights from Modern Greek and Turkish. *STUF—Language Typology and Universals* 68(3). 323–357.
- Banguoğlu, Tahsin. 2011. *Türkçenin grameri*. [Turkish grammar]. Ankara: Türk Dil Kurumu Yayınları.
- Bauer, Laurie. 2001. Compounding. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher & Wolfgang Raible (eds.), *Language typology and language universals*, 695–707. Berlin: Mouton de Gruyter.
- Bauer, Laurie. 2006. Compound. In K. Brown et al. (ed.), *Encyclopedia of language & linguistics*, 2nd edn. 719–726. London: Elsevier.
- Bauer, Laurie. 2011. Typology of compounds. In Rochelle Lieber & Pavol Štekauer (eds.), *The Oxford handbook of compounding*, 343–356. Oxford: Oxford University Press.
- Bisetto, Antonietta & Sergio Scalise. 2005. The classification of compounds. *Lingue e Linguaggio* 4(2). 319–332.
- Booij, Geert. 2005. *The grammar of words*. Oxford: Oxford University Press.
- Booij, Geert. 2010. *Construction morphology*. Oxford: Oxford University Press.

- Braun, Friederike & Geoffrey Haig. 2000. The noun/adjective distinction in Turkish: An empirical approach. In Aslı Göksel & Celia Kerslake (eds.), *Studies on Turkish and Turkic languages*, 85–92. Wiesbaden: Harrassowitz.
- Eraslan, Kemal. 2012. *Eski Uygur Türkçesi grameri*. [A grammar of Old Uighur]. Ankara: Türk Dil Kurumu Yayınları.
- Erdal, Marcel. 2004. *A grammar of Old Turkic*. Leiden: Brill.
- Ergin, Muharrem. 2011. *Türk dil bilgisi*. [Turkish grammar]. İstanbul: Bayrak Basım/Yayım/Tanıtım.
- Fabb, Nigel. 1998. Compounding. In Andrew Spencer & Arnold M Zwicky (eds.), *Handbook of morphology*, 66–83. Oxford: Blackwell.
- Forgács, Tamás. 2015. From Old Hungarian to Modern Hungarian. In Peter O Müller, Ingeborg Ohnheiser, Susan Olsen & Franz Rainer (eds.), *Word-formation: An international handbook of the languages of Europe*, vol. 3, 2079–2096. Berlin: Mouton de Gruyter.
- Göksel, Aslı. 2009. Compounds in Turkish. *Lingue e Linguaggio* 2. 213–236.
- Göksel, Aslı & Belma Haznedar. 2007. Remarks on compounding in Turkish. *MorboComp Project*, University of Bologna.
- Göksel, Aslı & Celia Kerslake. 2005. *Turkish: A comprehensive grammar*. London: Routledge.
- Güneş, Güliz. 2009. On the formation of [V + V] compounds in Turkish. Paper presented at the 2nd Mediterranean Graduate Students Meeting in Linguistics, Mersin, Turkey, 12–13 March 2009.
- Haspelmath, Martin. 2002. *Understanding morphology*. London: Arnold.
- Haspelmath, Martin. 2004. Coordinating constructions: An overview. In Martin Haspelmath (ed.), *Coordinating constructions*, 3–39. Amsterdam: Benjamins.
- Hatiboğlu, Vecihe. 1981. *Türk dilinde ikileme*. [Reduplication in Turkish]. Ankara: Ankara Üniversitesi Basımevi.
- Johanson, Lars. 1998. The structure of Turkic. In Lars Johanson & Éva Ágnes Csató (eds.), *The Turkic languages*, 30–66. London: Routledge.
- Kabak, Barış. 2007. Turkish suspended affixation. *Linguistics* 45(2). 311–347.
- Kageyama, Taro. 2010. Variation between endocentric and exocentric word structures. *Lingua* 120(10). 2405–2423.
- Kamali, Beste & Didem İkizoğlu. 2015. Compound stress in Turkish is phrase stress. In Deniz Zeyrek, Çiğdem Sağın-Şimşek, Ufuk Ataş & Jochen Rehbein (eds.), *Ankara papers in Turkish and Turkic linguistics*, 40–51. Wiesbaden: Harrassowitz Verlag.
- König, Wolf Dietrich. 1987. Nominal compounds in Turkish. *Zeitschrift für Sprachwissenschaft* 6(2). 165–185.
- Kornfilt, Jaklin. 1997. *Turkish*. London: Routledge.
- Kunduracı, Aysun & Aslı Göksel. 2016. Morphology: The base processor. In Jenny Audring, Francesca Masini & Wendy Sandler (eds.), *Quo vadis morphology? MMM10 Online proceedings*, 88–97. Leiden, Bologna & Haifa: University of Leiden, University of Bologna & University of Haifa.
- Larsson, Jenny H. 2002. Nominal compounds in the Baltic languages. *Transactions of the Philological Society* 100(2). 203–231.
- Lieber, Rochelle. 2005. English word-formation processes: Observations, issues, and thoughts on future research. In Pavol Štekauer & Rochelle Lieber (eds.), *Handbook of word-formation*, 375–427. Dordrecht: Springer.
- Lieber, Rochelle & Sergio Scalise. 2006. The lexical integrity hypothesis in a new theoretical universe. *Lingue e Linguaggio* 6(1). 7–32.

- Manolessou, Io & Symeon Tsolakidis. 2009. Greek coordinated compounds: Synchrony and diachrony. *Patras Working Papers in Linguistics* 1. 23–39.
- Marchand, Hans. 1969. *The categories and types of present-day English word-formation*. München: Beck.
- Masini, Francesca. 2006. Binomial constructions: Inheritance, specification and subregularities. *Lingue e Linguaggio* 5. 207–232.
- Navickaitė-Klišauskienė, Agnė. 2016. Latvian. In Peter O Müller, Ingeborg Ohnheiser, Susan Olsen & Franz Rainer (eds.), *Word-formation. An international handbook of the languages of Europe, vol. 5*, 3107–3124. Berlin: Mouton de Gruyter.
- Olsen, Susan. 2001. Copulative compounds: A closer look at the interface between syntax and morphology. In Geert Booij & Jaap van Marie (eds.), *Yearbook of morphology 2000*, 279–320. Dordrecht: Kluwer.
- Olsen, Susan. 2014. Coordinative structures in morphology. In Antonio Machicao Y Priemer, Andreas Nolda & Athina Sioupi (eds.), *Zwischen Kern und Peripherie: Untersuchungen zu Randbereichen in Sprache und Grammatik*, 269–286. Berlin: Mouton de Gruyter.
- Olsen, Susan. 2015. Composition. In Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen & Franz Rainer (eds.), *Word-formation: An international handbook of the languages of Europe*, 364–386. Berlin: Walter de Gruyter.
- Poppe, Nicholas. 1966. On some Turkic nominal phrases and compound nouns. *Reşid Rahmeti Arat için* [For Reşid Rahmeti Arat]. *Türk Kültürünü Araştırma Enstitüsü 19, Seri I, Sayı A2*, 372–379. Ankara: Ankara Üniversitesi Yayınevi.
- Renner, Vincent. 2008. On the semantics of English coordinate compounds. *English Studies* 89(5). 606–613.
- Scalise, Sergio & Antonietta Bisetto. 2009. The classification of compounds. In Rochelle Lieber & P. Štekauer (eds.), *The Oxford handbook of compounding*, 49–82. Oxford: Oxford University Press.
- Scalise, Sergio & Antonio Fábregas. 2010. The head in compounding. In Sergio Scalise & Irene Vogel (eds.), *Cross-disciplinary issues in compounding*, 109–126. Amsterdam: John Benjamins.
- Scalise, Sergio & Emiliano Guevara. 2006. Exocentric compounding in a typological framework. *Lingue e linguaggio* 2. 185–206.
- Scalise, Sergio & Irene Vogel. 2010. Why compounding? In Sergio Scalise & Irene Vogel (eds.), *Cross-disciplinary issues in compounding*, 1–20. Amsterdam: John Benjamins.
- Sertkaya, Osman Fikri. 2006. *Dede Korkut Kitabı – Dresden nüshasının giriş bölümü* [The book of Dede Korkut – Prologue of the Dresden copy]. İstanbul: Ötügen Neşriyat.
- Tekin, Talât. 2006. *Orhon yazıtları*. [Orkhon inscriptions]. Ankara: Türk Dil Kurumu Yayınlar.
- Ten Hacken, Pius. 2000. Derivation and compounding. In Geert Booij, Christian Lehmann & Joachim Mugdan (eds.), *Morphologie – Morphology: Ein Handbuch zur Flexion und Wortbildung – A handbook on inflection and word formation*, 349–360. Berlin: Mouton de Gruyter.
- Van Schaaik, Gerjan. 2002. *The noun in Turkish: Its argument structure and the compounding straitjacket*. Wiesbaden: Harrassowitz.
- Wälchli, Bernhard. 2005. *Co-compounds and natural coordination*. Oxford: Oxford University Press.