Morphosyntax of Two Turkish Subject Pronominal Paradigms[*]

Jeff Good and Alan C. L. Yu
University of Pittsburgh and University of Chicago

Abstract: Turkish exhibits two different sets of subject 'agreement markers' which show different morphosyntactic behavior from each other. It is argued here that one set of these markers are morphological suffixes while the other set are enclitics. This synchronic analysis is supported by diachronic facts which indicate that the agreement markers analyzed as suffixes have been suffixes throughout the reconstructible history of Turkic, while the agreement markers analyzed as clitics are more recent developments from reduced pronouns. A formal analysis of how these two sets of agreement markers are employed on Turkish verbs is developed within Head-driven Phrase Structure Grammar (HPSG).

0.   Introduction

Significant research has been done on the syntax and semantics of tense, mood, and aspect morphology in Turkish (see Sezer 2001 and references therein) but less has been done on the syntax of agreement. However, recent work reveals that Turkish agreement morphology is not as mundane as often assumed (cf. e.g., Orgun 1995, Good and Yu 2000, and Öztürk 2001). The present study provides an in-depth investigation on the divergent morphosyntactic behavior of two sets of subject markers in Turkish and is intended to further our understanding of the morphology and syntax of agreement in the language.

This paper begins with a presentation of a split in the morphosyntactic behavior of these two sets of subject markers in section 1. In section 2, we argue that this split results from that fact that one set of subject markers consists of post-lexical clitics while the other consists of lexical suffixes. In order to support our synchronic analysis, we will present a brief overview of the historical development of each set of subject markers in section 3. We will then present an account of verbal subject marking in Turkish in a lexicalist framework, namely Head-driven Phrase Structure Grammar (HPSG), in section 4. Some theoretical assumptions are clarified in section 5. A discussion of some of the implications of this study is given in section 6, and a brief conclusion appears in section 7.

1.   Pronominal Subject Endings in Turkish

The data used in this study is based on judgments of speakers of the Istanbul dialect of Turkish. The form and behavior of subject-marking paradigms can vary across dialects.

Turkish subject pronominal inflectional morphology employs four distinct suffixal paradigms. In this paper, we will concentrate on only two of these paradigms, leaving the imperative and the optative paradigms aside.[1] The two paradigms that we will focus on are given in (1).[2]

(1)		a.	k-*paradigm*		b.	z-*paradigm*	
		SINGULAR	PLURAL			SINGULAR	PLURAL
	1ST	-m	-k			-(y)Im	-(y)Iz
	2ND	-n	-nIz			-sIn	-sInIz
	3RD	-Ø	-Ø			-Ø	-Ø

The paradigm in (1a) (henceforth the *k*-paradigm after its first person plural from) only applies to verbal predicates that end with either the simple past suffix –*(y)DI*, shown in (2a), or the conditional suffix –*(y)sE*, shown in (2b).

(2) a. *dön-dü-**m***   b. *dön-se-**m***
    *dön-dü-**n***   *dön-se-**n***
    *dön-dü-**k***   *dön-se-**k***
    *dön-dü-**nüz***   *dön-se-**niz***
    turn-PAST-**PSN**   turn-COND-**PSN**

The paradigm in (1b) (henceforth the *z*-paradigm after its first person plural form) applies to all other predicates, both verbal and non-verbal (with the exception of the optative and imperative predicates mentioned above).

(3) a.   *gid-iyor-**uz***   'we are going'   **gid-iyor-k*
    b.   *adam-**ız***   'we are men'   **adam-k*
    c.   *iyi-**yiz***   'we are good   **iyi-k*
    d. * *git-ti-**yiz***   'we went'   *git-ti-**k***

The examples in (3a–c) show the types of predicates the *z*-paradigm can attach to. The corresponding ungrammatical forms with the *k*-paradigm endings are shown after each example. Example (3d) demonstrates that *z*-paradigm endings cannot affix to a verb in the simple past tense—the simple past suffix, like the conditional, can only take a *k*-paradigm ending.

    What has been presented so far are the most straightforward differences between the two paradigms. However, there are actually a number of idiosyncrasies in their behavior that demonstrate the existence of a major morphosyntactic distinction between them. We shall examine these in the immediately following sections.

1.2.   Ending Order Variation

In all the examples above, subject pronominal markers were the final ending on the predicate. This reflects their most common position throughout Turkish grammar. However, Sezer (1998) observes that when the predicate contains two or more tense, mood, and aspect (TMA) markers,

k-paradigm endings need not necessarily surface at the end of the predicate. For example, in (4), we show a partial paradigm of the verb *görmek* 'to see', illustrating that ordering variability can occur with *k*-paradigm personal endings without producing any difference in meaning. Specifically, the *k*-paradigm endings can appear predicate finally or between the two TMA markers. Such variability is not possible for the *z*-paradigm endings.

(4)   a.   *gör-dü-yse-**m***   *gör-dü-**m**-se*
            see-PAST-COND-**1SG**   see-PAST-**1SG**-COND
            'If I saw...'
       b.   *gör-dü-yse-**n***   *gör-dü-**n**-se*
            see-PAST-COND-**2SG**   see-PAST-**2SG**-COND
            'If you saw...'
       c.   *gör-dü-yse-**k***   *gör-dü-**k**-se*
            see-PAST-COND-**1PL**   see-PAST-**1PL**-COND
            'If we saw...'
       d.   *gör-dü-yse-**niz***   *gör-dü-**nüz**-se*
            see-PAST-COND-**2PL**   see-PAST-**2PL**-COND
            'If you (plural) saw...'

Variable ordering of the pronominal endings is not completely unconstrained however. The data in (5) shows that a pronominal ending must surface in verb-final position when the last two TMA markers employ conflicting pronominal paradigms. In this case, we have a combination of the conditional marker, which licenses only the *k*-paradigm endings, and the evidential marker, which licenses only the *z*-paradigm endings. Subject marking must be final, as seen in the grammatical (5a), but not the ungrammatical (5b).

(5)   a.   *oyna*   *-sa*   *-ymış*   *-**ız***
            play   COND   EVID   **1PL**
       b.* *oyna*   *-sa*   *-**k***   *-mış*
            play   COND   **1PL**   EVID
            'We were playing'

Judging from the data so far, one might assume that the reason that (5b) is ill-formed is merely due to the fact that the two TMA markers make use of different pronominal paradigms. However, as the data in (6) illustrate, even when both of the TMA markers are *z*-paradigm licensers, *z*-paradigm endings must still surface at the end of the predicate.

(6)   a.   *bul*   *-uyor*   *-**sun***
            find   PROG   **2SG**
            'You are finding'
       b.   *bul*   *-uyor*   *-muş*   *-**sun***
            find   PROG   EVID   **2SG**
       c. * *bul*   *-uyor*   *-**sun***   *-muş*
            find   PROG   **2SG**   EVID
            'You are apparently finding'

This inability of the *z*-paradigm endings to surface between TMA markers suggests that the distinction between the *k*- and the *z*-paradigms is more systematic than one might at first assume and goes beyond the relatively superficial differences of phonological shape and preceding TMA suffix.

It turns out that these two paradigms differ across a range of linguistic parameters which allows for a more principled explanation than merely attributing such morphological idiosyncrasies to chance. In the remainder of this paper, we will explicate the nature of this bipartite behavior on both synchronic and historical grounds. We will also present a formal account that attempts to succinctly capture the different behavior of each paradigm.

2. Clitics vs. Lexical Affix-hood of the Turkish Subject Pronominal Endings

The difference between the *k*- and the *z*-paradigms, as we shall argue, is a matter of their formal status in the lexicon. That is, we claim that the *k*-paradigm endings are lexical suffixes but the *z*-paradigm endings are post-lexical clitics.

To support this claim, we rely on the diagnostic conditions that were provided in Zwicky and Pullum (1983). Their criteria are reproduced in (7A-F).

(7) A.   Clitics can exhibit a low degree of selection with respect to their hosts, while affixes exhibit a high degree of selection with respect to their stems.
    B.   Arbitrary gaps in the set of combinations are more characteristic of affixed words than of clitic groups.
    C.   Morphological idiosyncrasies are more characteristic of affixed words than of clitic groups.
    D.   Semantic idiosyncrasies are more characteristic of affixed words than of clitic groups.
    E.   Syntactic rules can affect affixed words, but cannot affect clitic groups.
    F.   Clitics can attach to material already containing clitics, but affixes cannot.

Not all of these conditions can be applied to our data. However, three of the seven do, and each of those three indicates that *k*-paradigm endings are suffixes and *z*-paradigm endings are clitics. We have already seen the data that has bearing on (7A). The *k*-endings only follow two verbal suffixes while the *z*-endings follow all other verbal suffixes as well as non-verbal predicates. The variable ordering of *k*-paradigm endings is fairly idiosyncratic in Turkish grammar as it is the only case where subject marking is not at the very end of the sentence. We know of no comparable idiosyncratic behavior for *z*-endings. So, criterion (7C) also favors our claim. A conjunction reduction process in Turkish, known as suspended affixation, to be illustrated below in section 2.2, treats verbs plus *k*-endings as whole constituents whereas it does not treat the combination of verb + *z*-paradigm ending as a constituent. Thus, also by criterion (7E), *k*-endings behave like suffixes and *z*-endings like clitics.

2.1. Phonological Evidence

Zwicky and Pullum also point out that clitics are generally accentually dependent—that is, they are not usually stressed. Default Turkish stress is word-final. The one syllabic *k*-paradigm ending, -*nIz* '2PL', is stressed when it is word-final, whereas *z*-paradigm endings are never stressed, as exemplified in (8). Thus, word-final *k*-paradigm endings behave as though they are truly part of the word they suffix to, with respect to stress assignment, while *z*-paradigm endings do not, taking on an additional common property of clitics.

(8) *z*-paradigm ending   *k*-paradigm ending
     unut   -a'cak -sınız   *unut   -tu   -'nuz*
     forget FUT   **2PL**   forget   PAST   **2PL**
     'you will forget'   'you forgot'

Looking at the data from Turkish in light of the criteria set forth by Zwicky and Pullum is part of the justification of our claim. However, as pointed out by Miller (1992), the most definitive test for clitic-hood is the ability of the clitic to participate in coordination—that is, the possibility of its having wide scope over a conjunction of hosts. This, as we shall show, is allowed for by the *z*-paradigm, but not the *k*-paradigm.

2.2. Suspended Affixation

The evidence for the coordination facts mentioned above can be found in what linguists who work on Turkic languages refer to as suspended affixation (Lewis 1967, Orgun 1995). This is a phenomenon in Turkish where endings are optionally omitted from all conjuncts in a coordinated structure except for the last one. The endings on the last conjunct then have semantic scope over all the conjuncts. Some examples are given in (9). We should be quick to point out that the facts of suspended affixation are much broader than can be adequately discussed and accounted for here. We will only concentrate here on the data that is relevant to the morphosyntactic distribution of the subject pronominal endings.

(9) a. [*genç   ve   büyüğ*]-***üm***
        [young   and   big]-**1SG**
    b. [*genç*]-***im***   ve   [*büyüğ*]-***üm***
        [young]-**1SG**   and   [big]-**1SG**
        'I am young and big'
    c. [*hastane-ye   gid-iyor,   o-nu   gör-üyor*]-***sunuz***
        [hospital-DAT   go-PROG   3SG-ACC   see-PROG]-**2pl**
    d. [*hastane-ye   gid-iyor*]-***sunuz***,   [*o-nu   gör-üyor*]-***sunuz***
        [hospital-DAT   go-PROG]-**2PL**   [3SG-ACC   see-PROG]-**2PL**
        'You all are going to the hospital and seeing him/her.'

The example in (9a) illustrates that the first person singular ending –*(y)Im* can have semantic scope over both the predicate 'young' and the predicate 'big'. The unsuspended counterpart of (9a) is given in (9b). The first person plural ending –*(y)Iz* in (9c) has scope over both *gidiyor* 'go' and *görüyor* 'see'. The unsuspended version of (9c) is given in (9d). These examples illustrate that suspended affixation can operate for *z*-paradigm endings in both verbal and non-verbal predicates.

The application of this suspended affixation construction, however, is not totally unrestricted. Endings from the *k*-paradigm are prohibited from participating in the construction, as illustrated in (10).

(10) a. [*ev-e*   *gel-ir,*   *sana*   *yardım*   *ed-er*]-**iz**
          [home-DAT   come-AOR   you-DAT   help   do-AOR]-**1PL**
          'We'll come home and then we'll help you.'
      b.* [*ev-e*   *gel-di,*   *sana*   *yardım*   *et-ti*]-***k***
          [home-DAT   come-PAST   you-DAT   help   do-PAST]-**1PL**
          'We came home and then we helped you.'
      c. [[*ev-e*   *gel-di-**k***],   [*sana*   *yardım*   *et-ti-**k***]]
          [[home-DAT   come-PAST-**1PL**]   [you-DAT   help   do-PAST-**1PL**]]
          'We came home and then we helped you.'

Suspended affixation is observed in (10a) since the subject ending required is from the *z*-paradigm. The sentence in (10b), on the other hand, is ill-formed since the subject pronominal ending suspended is from the *k*-paradigm. The *k*-paradigm ending is required there since the simple past tense suffix –*(y)DI* only takes endings from this paradigm. The well-formed equivalent of (10b) is given in (10c). The fact that suspended affixation is impossible for the *k*-paradigm endings suggests there is strong lexical affinity of the *k*-paradigm endings to the simple verbs ending with the past tense suffix –*(y)DI*, which would be expected of true suffixes. These same basic facts for suspended affixation are true for the conditional suffix –*(y)sE*, the other suffix taking the *k*-paradigm.

2.3   Conclusion

Using primarily atheoretical criteria as our diagnostics, the claim has been made that the *k*-paradigm endings in Turkish are lexical affixes while the *z*-paradigm endings are post-lexical clitics. Given that the terms *affix* and *clitic* are often used in subtly, but importantly, different ways in various theories, it would be worthwhile for us to summarize the basic argument being put forth here in light of potential confusion caused by the different possible senses of these terms.

It is clear that *k*-paradigm endings and *z*-paradigm endings in Turkish behave systematically different. On the one hand, *k*-paradigm endings can be stressed, show positional variability (and, thereby, can appear internally in the Turkish verb complex), only appear after two verbal suffixes, and cannot have scope over multiple, conjoined verbs. On the other hand, *z*-paradigm endings are never stressed, always occur at the right-edge of their host, can appear after verbal suffixes and non-verbal predicates, and can have scope over multiple conjoined verbs.

These differences in the two paradigms point to an analysis across phonological, morphological, and syntactic dimensions where the *k*-paradigm endings should be treated as affixes and the *z*-paradigm endings as clitics, as the terms have been traditionally understood—that is, where affixes are taken to be morphological entities and clitics are taken to be phonologically-dependent syntactic entities. We do not take this to mean that these two sets of endings will necessarily fall on different sides of the divide between the labels 'affix' and 'clitic' in every linguistic theory, however. For example, work like that of Halpern (1994: 101) proposes the existence of a category between affixes and clitics—lexical clitics. We are unaware of any data from Turkish which indicates the existence of such morphemes in the language. So, we have not addressed the possibility that either the *k*-paradigm or *z*-paradigm endings in Turkish are of this class.

It is not inconceivable that another researcher would want to use the label 'lexical clitic' for either the *k*- or *z*-paradigm endings. However, no matter what formal theory one chooses to adopt, all the data uncovered so far unambiguously shows a two-way split where *k*-paradigm endings are basically 'affixal' and *z*-paradigm endings are basically 'clitic-like'. We take this to mean that the most straightforward analysis of the split between the two paradigms is one where the *k*-paradigm endings are treated as morphological elements and the *z*-paradigm endings are treated as syntactic elements, and this will be clearly reflected in our formal analysis in section 4. Furthermore, because the data demonstrates such a 'clean' split in the behavior of the two paradigms, we suspect any successful analysis of the Turkish facts seen here will have to make use of some sort of morphology/syntax distinction along the lines of the one we will make.

3. Historical Perspective

So far, we have seen ample synchronic evidence that points to a morphological dichotomy between the *k*- and the *z*-paradigms. In this section, we shall further examine this clitic vs. suffix distinction by discussing the apparent historical origin of these two paradigms.

Old Turkic, the oldest attested Turkic language, did not have suffixing subject agreement markers. Sentences with pronominal subjects were formed by putting a pronoun at the end of the predicate, as seen in (11a) and (11b). The contrasting example in (11c) does not have a pronominal subject.

(11) a. *kel-ür*   ***ben***   (Adamović 1985: 27)
         come-AOR   **1SG**
         'I'm coming.'
     b. *sen*   *kisi*   ***sen***
         2SG   man   **2SG**
         'You are a man.'
     c. *altun*   *sari*
         gold   yellow
         'Gold is yellow.'

Cliticized forms of the predicate-final pronouns appear in thirteenth century texts. A typical paradigm from that period is given in (12). As can be seen in (12), third-person subjects did not

make use of cliticized pronouns. Rather they used a different morphological suffix, which did not develop into a subject-marking suffix.

(12) Singular *(ben)* *bay-**van*** 'I am rich.' (Adamović 1985: 27)
(sen) *bay-**sin*** 'You are rich.'
(ol) *bay-**durur*** 'He/she/it is rich.'
Plural *(biz)* *bay-**uz*** 'We are rich.'
(siz) *bay-**siz*** 'You (PL) are rich.'
(anlar) *bay-**durur(lar)*** 'They are rich.'

After several sound changes and morphological shifts, the cliticized versions of the pronouns seen in Old Turkic evolved into the *z*-paradigm in Modern Turkish.

    Although the historical development of the *z*-paradigm is rather transparent, the historical origin of the *k*-paradigm is still somewhat obscure. Shaw (1877) points out that the Old Turkic preterite was formed via the possessive construction as in (13).

(13) *qil-d-um* (Adamović 1985: 184)
do-NOM-1SG.POSS
'My action of doing (exists)'

The morpheme *-d-* in (13) serves as a nominalizing suffix while the *-um* suffix is the first person singular possessive suffix. According to Adamović (1984), a reference in Kaşgarli (1939: 60–63) mentions that the preterite was in mutually exchangeable relations with the periphrastic formed by the deverbal nominal suffix *-duq̱/-dük* in Oghus, Suwar and probably Kipchak in the eleventh century, giving us further insight into the nominal nature of the preterite form. This is exemplified in (14).

(14) a. *ben* *baq-duq* ~ *baq-d-um* (Adamović 1985: 185)
    1SG look-NOM look-NOM-1ˑSG.POSS
b. *sen* *baq-duq* ~ *baq-d-uŋ*
    2SG look-NOM look-NOM-2SG.POSS
c. *ol* *baq-duq* ~ *baq-d-i*
    3SG look-NOM look-NOM-3SG.POSS
d. *biz* *baq-duq* ~ *baq-d-umuz*
    1PL look-NOM look-NOM-1PL.POSS
e. *siz* *baq-duq* ~ *baq-d-uŋuz*
    2PL look-NOM look-NOM-2PL.POSS
f. *anlar* *baq-duq* ~ *baq-d-iler*
    3SG look-NOM look-NOM-3PL.POSS

The *-dV-* found in the possessed forms in (14) appears to have been later reanalyzed as a past tense marker and the former possessive suffixes of the non-periphrastic preterite construction were treated as the subject-marking suffixes for predicates that ended in the new past tense morpheme. This change is schematized in (15).[3]

(15)   *Before reinterpretation*   *After reinterpretation*

a.   *baq-d-um*   > *baq-du-m*
    look-NOM-1˙SG.POSS   look-PAST-1SG

b.   *baq-d-uŋ*   > *baq-du-ŋ*
    look-NOM-2SG.POSS   look-PAST-2SG

c.   *baq-d-i*   > *baq-di-Ø*
    look-NOM-3SG.POSS   look-PAST-3SG

d.   *baq-d-umuz*   > *baq-du-muz*
    look-NOM-1PL.POSS   look-PAST-1PL

e.   *baq-d-uŋuz*   > *baq-du-ŋuz*
    look-NOM-2PL.POSS   look-PAST-2PL

f.   *baq-d-iler*   > *baq-di-ler*
    look-NOM-3PL.POSS   look-PAST-3PL

Assuming the historical development of the *k*-paradigm from the preterite is valid, the application of the *k*-paradigm to the predicates marked with the conditional *-(y)sE* appears to be a case of analogical extension at a later stage in Turkic. This leveling would presumably have been motivated by the fact that the conditional suffix, like the past suffix, ends in a vowel.

What this historical scenario illustrates is that the *k*-paradigm endings appear to never have been independent lexical items at any stage of the traceable history of the Turkic language, as opposed to the *z*-paradigm endings, which originate from full independent words that were reduced, in several stages. This separate historical development offers an account of why the split of the *k*-paradigm endings as suffixes and the *z*-paradigm endings as clitics exists in today's language.

While we now understand the origins of the two paradigms, we have yet to understand how variable affix ordering developed for endings in the *k*-paradigm, which is of potential interest since it is one of the more idiosyncratic properties of that paradigm. To understand the most likely historical source of this variable ordering, it is useful to look at certain periphrastic tense constructions in Turkish which we exemplify using the pluperfect in (16). Historically, the pluperfect tense was formed by the past tense of a verb followed by the past tense of the verb *i* 'to be', and this construction is still available today. As illustrated in (16), such periphrastic constructions allow the personal endings to be marked on either the matrix verb or the verb 'to be'. (The third-person plural forms seen in (16f) and (17f) make use of the plural suffix, as opposed to a true person suffix. We discuss the ordering properties of this suffix in Modern Turkish section 4.4.)

(16) Two possible ways of realizing the periphrastic pluperfect (Lewis 1953: 92)

a.   *git-ti i-di-**m***   *git-ti-**m** i-di*   'I had gone'
b.   *git-ti i-di-**n***   *git-ti-**n** i-di*   'You had gone'
c.   *git-ti i-di-**Ø***   *git-ti-**Ø** i-di*   'He had gone'
d.   *git-ti i-di-**k***   *git-ti-**k** i-di*   'We had gone'
e.   *git-ti i-di-**niz***   *git-ti-**niz** i-di*   'You had gone'
f.   *git-ti i-di-**ler***   *git-ti-**ler** i-di*   'They had gone'

The person-marking variation in (16) is presumably attributable to several factors. First, both the main verb and the *i* auxiliary verb end in a *k*-paradigm taking suffix. So, there is no conflict in

realizing the *k*-paradigm affixes on either verb with respect to morphological subcategorization. The past-tense form of the matrix verb, also, exists independently as a possible inflected past tense verb in the language. Thus, all of the person-marked verbs in (16) would be found elsewhere in the language.

Moreover, as Haiman (1977) observed for Turkish, the null third-person marker creates an ambiguous context where the marker could be interpreted as either marked on the first or the second verb for instances of the construction, like those in (16c), with third-person subjects. In accordance with Watkins' Law (Haiman 1977), such ambiguity could have contributed to the rise of the variability in the forms where subject marking on the verb is overt.

The Turkish verb forms with multiple TMA suffixes taking *k*-paradigm endings resulted from the reduction of the auxiliary verb *i* seen in (16). To explain the affix ordering variability in forms like those in (4), then, we simply need to assume that this reduction was not accompanied by a loss in the multiple possibilities for the position of person marking. Thus, forms like those given in (17) can be understood as the expected development from the reduction of the auxiliary seen in the forms in (16).

(17) Grammaticalized pluperfect formation (Lewis 1953: 92)
a.	*git-ti-ydi-**m***	*git-ti-**m**-di*	'I had gone'
b.	*git-ti-ydi-**n***	*git-ti-**n**-di*	'You had gone'
c.	*git-ti-ydi-**Ø***	*git-ti-**Ø**-ydi*	'He had gone'
d.	*git-ti-ydi-**k***	*git-ti-**k**-ti*	'We had gone'
e.	*git-ti-ydi-**niz***	*git-ti-**niz**-di*	'You had gone'
f.	*git-ti-ydi-**ler***	*git-ti-**ler**-di*	'They had gone'

Essentially the same scenario can account for ordering variability involving the conditional suffix, which also historically entered into periphrastic verb construction with the verb *i* 'to be'.

4.   Analysis

In the remainder of this paper, we will briefly sketch out a formal account of the data. However, before we move into the discussion, we shall review the generalizations that we intend to capture in formal terms. They are summarized in (18).

(18) i.   *Stress assignment*: unlike *k*-paradigm endings, *z*-paradigm endings cannot receive stress.

ii.   *Morphological selectivity*: *k*-paradigm endings can only suffix to verbal predicates of certain TMA categories while *z*-paradigm endings can also suffix to non-verbal predicates.

iii.   *Variable ordering*: *z*-paradigm endings can only occur word-finally while *k*-paradigm endings can also occur word-internally without a change in meaning.

iv.   *Wide scope in coordination*: *z*-paradigm endings can have wide scope over more than one conjunct in coordination, while *k*-paradigm endings cannot.
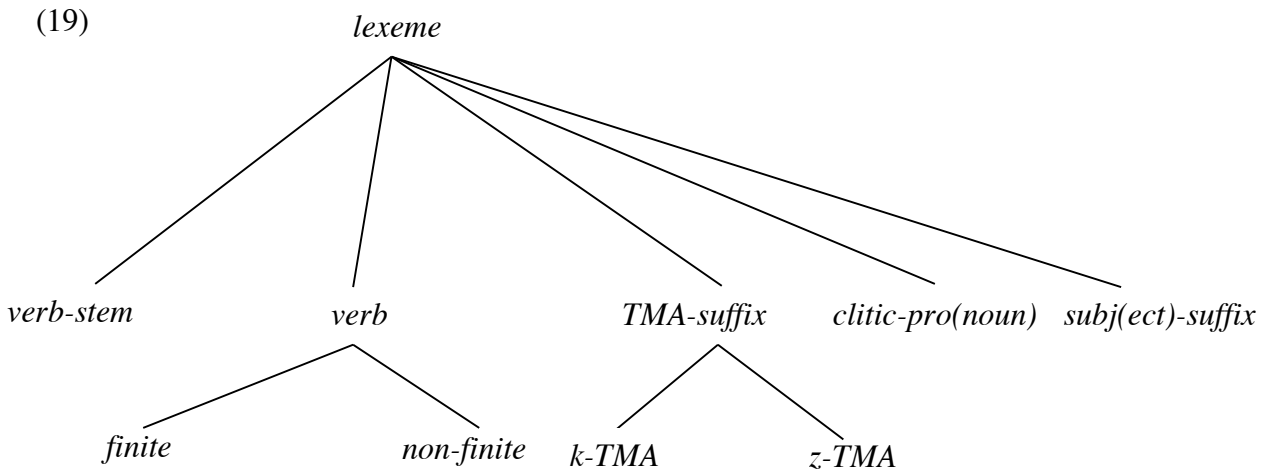
We will use the lexicalist framework of Head-driven Phrase Structure Grammar (HPSG) in formalizing the data. We follow the conventions of Sag and Wasow (2000) when defining

syntactic constructions and adapt ideas found in Koenig (1999) when defining morphological ones.

4.1.   Morphological Realization of the k-*paradigm endings*

The critical elements we adopt from Koenig (1999) are online morphological type constructions (1999: 51–83) and complex morphological signs (1999: 88–134). We diverge from certain formal details found in Koenig (1999) in the interests of exposition.

The tree in (19) illustrates the hierarchy of the lexical types we will need to assume for Turkish in our analysis.

(19)

```
                          lexeme


  verb-stem        verb              TMA-suffix     clitic-pro(noun)   subj(ect)-suffix


        finite          non-finite   k-TMA              z-TMA
```

The type hierarchy in (15) states the existence of five different major types of lexemes which will be used in the analysis: *verb-stem*, *verb*, *TMA-suffix*, *clitic-pro*, and *subj-suffix*. These lexemes types are roughly analogous to the traditional idea of part of speech. However, Koenig's framework allows for a much more richly articulated set of lexical categories than simply using the traditional categories of noun, verb, etc. The five major types of lexemes posited in (19) are not intended to be exhaustive for Turkish. They simply represent the types which will be referred to in the analysis to be presented below.

The first of the five types is *verb-stem*. This is simply the class of lexical items in Turkish which can take verbal TMA suffixes. It includes both verb roots (which can serve as verb stems on their own) and verb roots which have been suffixed with various derivational morphemes. The TMA suffixes which can attach to *verb-stems* are of the type *TMA-suffix*, and this type is subdivided into *k-TMA* suffixes (which take *k*-paradigm endings) and *z-TMA* suffixes (which take *z*-paradigm endings). The third type of lexeme is the *verb*. We will formalize lexemes of this type as consisting of a *verb-stem* with one or more TMA suffixes attached to it as in (20). The final two types of lexemes are the subject markers, of which there are two types: the clitic *z*-paradigm markers (*clitic-pro*) and the suffix *k*-paradigm markers (*subj-suffix*) (both of which are discussed in more detail below).

(20)

$$
\begin{bmatrix}
verb \\
\text{PHON} \quad \begin{bmatrix} \boxed{2} \oplus \boxed{1} \end{bmatrix} \\
\\
\mu\text{-STRUC} \quad \begin{bmatrix} \text{HEAD-DGHTR} & \begin{bmatrix} \textit{TMA-suffix} \\ \text{PHON} \quad \boxed{1} \end{bmatrix} \\ \\ \text{NON-HEAD-DGHTRS} & \left\langle \begin{bmatrix} \textit{verb-stem} \mid \textit{verb} \\ \text{PHON} \quad \boxed{2} \end{bmatrix} \right\rangle \end{bmatrix}
\end{bmatrix}
$$

As mentioned above, the matrix in (20) represents the morphological construction for the type *verb*. The basic intuition behind (20) is that a verb in Turkish consists of a verb stem plus at least one verbal TMA suffix. (The only instance where verb stems can surface without some suffix in Turkish is the second person singular imperative which we take to require a separate, unformalized construction.)

There are several dimensions to the formalization in (20) worth pointing out in more detail. First, it states that a lexeme of type *verb* has a morphological structure (*μ-struc* in (20)) consisting of a *verb* or a *verb-stem* to which a lexeme of the type *TMA-suffix* is attached. The *TMA-suffix* is viewed as the head of this morphological construction since it is what determines the overall morphosyntactic behavior of the verb (for example, whether it is a *k*-paradigm-taking or a *z*-paradigm-taking verb). The morphological constituent the *TMA-suffix* attaches to is taken to be a morphological daughter in the construction. Such a formalization can be understood as corresponding to the idea that the *TMA-suffix* subcategorizes for its host. The *TMA-suffix* is specified as allowing its daughter to be either of type *verb-stem* or of type *verb* (the "|" symbol is used to represent the logical operator "or"). This straightforwardly allows multiple suffixes to appear on the same verb, which is seen in, for example, the data in (4) and (5a). (In (21) the structure of a verb with multiple verbal suffixes is elaborated.) The first TMA suffix attaches to an element of type *verb-stem* to create an element of type *verb* which itself can then appear with further suffixes.

Another important aspect of the formalization in (20) further relates to the fact that the complex morphological structure is specified as being of type *verb*. While the category of *verb-stem* is a purely morphological category, the category of *verb* is taken to be both morphological and syntactic. This will be important in later discussion. We should make it clear, at this point, however, that, just because some constituent is of type *verb*, this does not necessarily mean it can be a surfacing syntactic constituent—the syntax might impose requirements on verbs beyond simple lexical category specification.

The final aspect of (20) which needs to be discussed is the PHON attribute. This is a formalization of the phonological realization of the construction. Specifically, it states that the phonology of the morphologically complex constituent consists of the phonology of the verb or verb stem (represented by the boxed "2") concatenated with the phonology of the suffix (represented by the boxed "1"). This aspect of (20) is simply reflects the fact that elements of type *TMA-suffix* are phonologically suffixing. Since the label for the category of *TMA-suffix* is simply a mnemonic for its actual use, the fact that it is suffixing needs to be explicitly formalized.

An example of the application of (20) is given in (21) which is based on the stem of the verb *gördüysem* 'If I saw', seen in (4a). This example does not include the subject marking since that is taken care of by constructions developed below in (22) and (23). The construction in (20) is instantiated twice in (21) since the verb contains two affixes of the type *TMA-suffix*—for this particular verb the particular *TMA-suffixes* are both *k-TMA* suffixes. The construction in (20) states that the head of the morphological construction must be of some element of the general type *TMA-suffix*. So, for any one particular instantiation of the construction, the head will be either one of the two subtypes of TMA suffixes, a *k-TMA* or a *z-TMA* suffix.

(21)

$$
\begin{bmatrix}
verb \\
\text{PHON} \quad \left[\boxed{4} \oplus \boxed{1}\right] \\
\mu\text{-STRUC} \quad
\begin{bmatrix}
\text{HEAD-DGHTR} \quad
\begin{bmatrix}
k\text{-}TMA \\
\text{PHON} \quad \boxed{1}\left[yse \text{ 'COND'}\right]
\end{bmatrix} \\
\text{NON-HEAD-DGHTRS} \quad
\begin{bmatrix}
verb \\
\text{PHON} \quad \boxed{4}\left[\boxed{3} \oplus \boxed{2}\right] \\
\mu\text{-STRUC} \quad
\begin{bmatrix}
\text{HEAD-DGHTR} \quad
\begin{bmatrix}
k\text{-}TMA \\
\text{PHON} \quad \boxed{2}\left[dü \text{ 'PAST'}\right]
\end{bmatrix} \\
\text{NON-HEAD-DGHTRS} \quad
\begin{bmatrix}
verb\text{-}stem \\
\text{PHON} \quad \boxed{3}\left[gör \text{ 'see'}\right]
\end{bmatrix}
\end{bmatrix}
\end{bmatrix}
\end{bmatrix}
\end{bmatrix}
$$
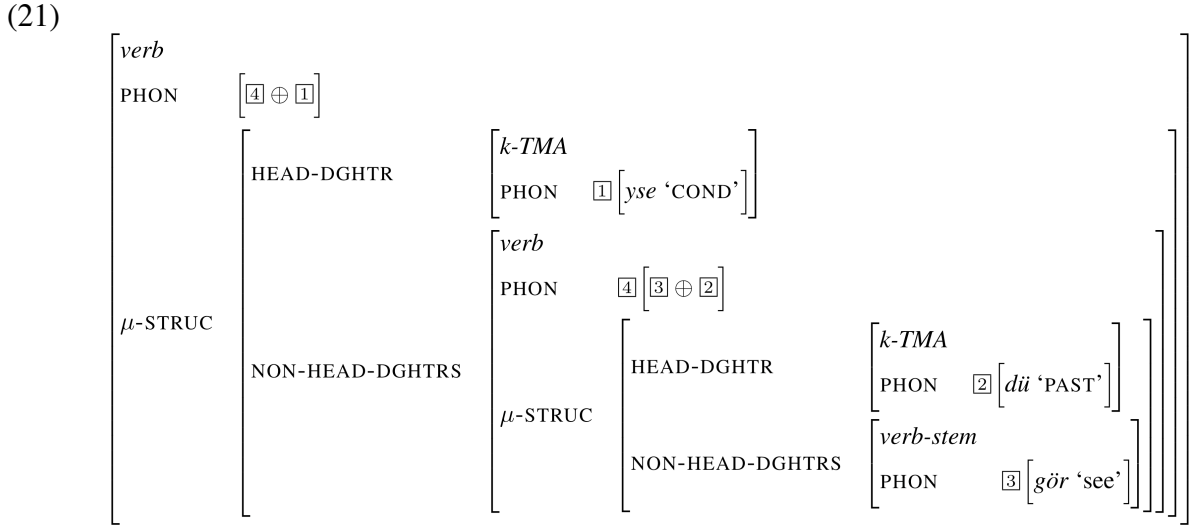
In (21) we see that the verb *gördüysem* is formalized at the top level as an instantiation of (20) with its two morphological components being the daughter verb *gördü* and the verbal suffix *yse*. Furthermore, the daughter verb *gördü* is itself an embedded instantiation of (20) with morphological components consisting of the verb stem *gör* and the verbal suffix *dü*.

Returning to the type hierarchy in (19), it further defines two types of verbs, *finite* and *non-finite*. The former are verbs which take *k*-paradigm endings, and, therefore, show morphological subject markers. The latter are those verbs which take the *z*-paradigm clitic endings. Since *z*-paradigm endings are taken, here, to be added in the syntax, these verbs have been labeled *non-finite* to indicate that, morphologically, they do not take subject markers.

Whether a verb is of lexical type *finite* or *non-finite* is determined by its outermost verbal suffix. Within the present formalization, this is the suffix added by the outermost instantiation of the construction in (20) in forming a particular verb. For example, the outermost instantiation of the construction defined in (20) for the verb formalized in (21) is the one where the *k-TMA* suffix *yse* is added to the verb, making it *finite*. We formalize this aspect of the analysis with the two constructions given in (22). The construction in (22a) simply states that any morphological constituent whose head is of type *k-TMA* (a *k*-paradigm-taking suffix) will be of type *finite*. The construction in (22b) states that any constituent ending in a *z-TMA* (a *z*-paradigm-taking suffix) will be of type *non-finite*. In addition, (22b) stipulates that a *z-TMA* suffix must have as its morphological daughter a verb or verb stem which has an unfilled SUBJ element (indicated by the empty square brackets within the angled brackets)—that is, it can only suffix to verbs which are not already specified for the subject. We will see shortly why this is necessary for the broader analysis.

(22) a.

$$
\begin{bmatrix}
\textit{finite} \\[4pt]
\mu\text{-STRUC} \quad \begin{bmatrix} \text{HEAD-DGHTR} & \begin{bmatrix} \textit{k-TMA} \end{bmatrix} \end{bmatrix}
\end{bmatrix}
$$

b.

$$
\begin{bmatrix}
\textit{non-finite} \\[4pt]
\text{SUBJ} \quad \langle [\ ] \rangle \\[6pt]
\mu\text{-STRUC} \quad
\begin{bmatrix}
\text{HEAD-DGHTR} & \begin{bmatrix} \textit{z-TMA} \end{bmatrix} \\[8pt]
\text{NON-HEAD-DGHTRS} & \left\langle \begin{bmatrix} \textit{verb-stem} \mid \textit{verb} \\ \text{SUBJ} \quad \langle [\ ] \rangle \end{bmatrix} \right\rangle
\end{bmatrix}
\end{bmatrix}
$$

We further give the construction in (23) in order to formalize the fact that *finite* verbs (that is, those that end in a *k*-paradigm-taking suffix) can be marked for subject agreement.

(23)

$$
\begin{bmatrix}
\textit{finite} \\[4pt]
\text{PHON} \quad \begin{bmatrix} \boxed{1} \oplus \boxed{2} \end{bmatrix} \\[4pt]
\text{SUBJ} \quad \langle\ \rangle \\[6pt]
\mu\text{-STRUC} \quad
\begin{bmatrix}
\text{HEAD-DGHTR} & \begin{bmatrix} \textit{finite} \\ \text{PHON} \quad \boxed{1} \\ \text{SUBJ} \quad \langle \boxed{3} \rangle \end{bmatrix} \\[12pt]
\text{NON-HEAD-DGHTRS} & \left\langle \boxed{3} \begin{bmatrix} \textit{subj-suffix} \\ \text{PHON} \quad \boxed{2} \end{bmatrix} \right\rangle
\end{bmatrix}
\end{bmatrix}
$$

The construction in (23) has both important morphological and syntactic dimensions. Morphologically, it says that *finite* verbs can have an element of type *subj-suffix* (a suffix from the *k*-paradigm) as a daughter. Syntactically, the construction says that an element of type *subj-suffix* can fill a verb's SUBJ slot—that is, it can act as the subject of the verb. When this is the case, the SUBJ slot of the verb becomes empty (as indicated by the empty angle brackets at the top of the construction).

The addition of a *subj-suffix* to a finite verb is not taken to affect the morphological category of a verb. Thus, the constituent defined by the construction in (23) is still of type *finite*. This is important for the analysis, given data like that in (4), since it means that nothing prevents the addition of another *k-TMA* marker on a verb which already shows subject agreement. The fact

that this construction causes the SUBJ slot of the verb to be filled, however, means that, once it is instantiated, no *z-TMA* marker can attach to the verb because, as seen in (22b), these suffixes are stipulated as only appearing on verbs or verb stems where the SUBJ slot is not filled. This makes this analysis consistent with facts reported in Good and Yu (1999). It is also consistent with the ungrammaticality of the verb in (5b).

To complete the analysis, we will need to ensure that one additional requirement is met for verbs of type *finite*. Morphologically, they can be allowed to have a non-filled subject slot—this is, of course, necessary in order for them to be constructed in the first place, via a series of multiple suffixations, only one of which can actually add subject suffix. However, the construction in (23) must be instantiated at some point in the construction of a verb of type *finite* in order for it to be *syntactically* well-formed. This will force (sometimes null) subject agreement to be found on finite verbs before they enter into any syntactic constructions. We take this restriction to be one which the syntax, but not the morphology, imposes on verbs of type *finite*. Verbs of type *finite* are both possible morphological and syntactic constituents, but nothing prevents the morphology and the syntax from each putting different stipulations on their form.

In addition to capturing the basic morphosyntactic facts of verbs marked with *k*-paradigm endings, this formulation of the nature of the subject agreement found on them has two important other consequences. First, it prohibits the possibility that these endings could participate in suspended affixation, discussed in section 2.2, since it implies that every finite verb will appear with subject agreement before it can enter into any phrasal construction, making it impossible for the subject suffix of one verb to also have scope over another verb requiring a subject suffix. Thus, (23) correctly formalizes the data seen in (10b). Second, by formalizing elements of type *subj-suffix* (i.e., *k*-paradigm endings) as fully morphologically integrated into their verbs, this analysis correctly predicts they should take stress when they are the word-final syllable, following the general pattern of word-final stress in Turkish.

4.2.   *Syntactic realization of the z-paradigm*

Regarding the treatment for the *z*-paradigm endings, recall that we take the data we have found to indicate that *z*-paradigm endings are realized phrasally. To formalize this, we propose to treat the *z*-enclitics as independent signs that are combined with verbal or nominal predicates through mechanisms in the syntax. The signs that characterize the *z*-enclitics are taken to be exemplified by (19), which gives the specific features of the first person plural *z*-paradigm ending *-(y)Iz*. This pronominal enclitic sign has three crucial specifications: Its phonological form, its person features, and its number feature.

At this point in the analysis, we are dealing with syntactic, not morphological, generalizations. Therefore, our formal conventions, though largely the same as above, are taken not from Koenig (1999), but, rather, are adapted from Sag and Wasow (2000). The two works can be straightforwardly used together as they are formal treatments of different aspects of grammar (morphology and syntax, respectively) within comparable lexicalist models.

(24)

$$
\begin{bmatrix}
\textit{clitic-pro} \\[4pt]
\text{PHON} \quad \begin{bmatrix} \textit{(y)Iz} \end{bmatrix} \\[8pt]
\text{SYNSEM} \quad \begin{bmatrix} \text{HEAD} \begin{bmatrix} \text{AGR} \begin{bmatrix} \text{PER} & 1 \\ \text{NUM} & \text{pl} \end{bmatrix} \end{bmatrix} \end{bmatrix}
\end{bmatrix}
$$

Not formalized here is the fact the elements of the type *clitic-pro* cannot be phonologically stressed. Doing so would require a working out formal details of phonological representations in Turkish which are outside the scope of this paper. The construction below in (25) will formalize the fact that the clitic follows the verb. The terms SYNSEM and HEAD in (24) have particular interpretations within HPSG. Here, all that is critical about them is the fact that person and number information is stored within them as part of the feature AGR (agreement).

A sign like the one in (24) can combine with any predicate that requires a subject—thus, it can apply to verbs which are of type *non-finite*. In order to demonstrate how such combination works, it is necessary to formalize how a given predicate can specify restrictions on its subject in the syntax. Recall that the construction in (23) specifies how certain predicates can fulfill their subject requirements morphologically. In (25) we give a construction which states restrictions on how *non-finite* predicates can fulfill their subject requirements syntactically.
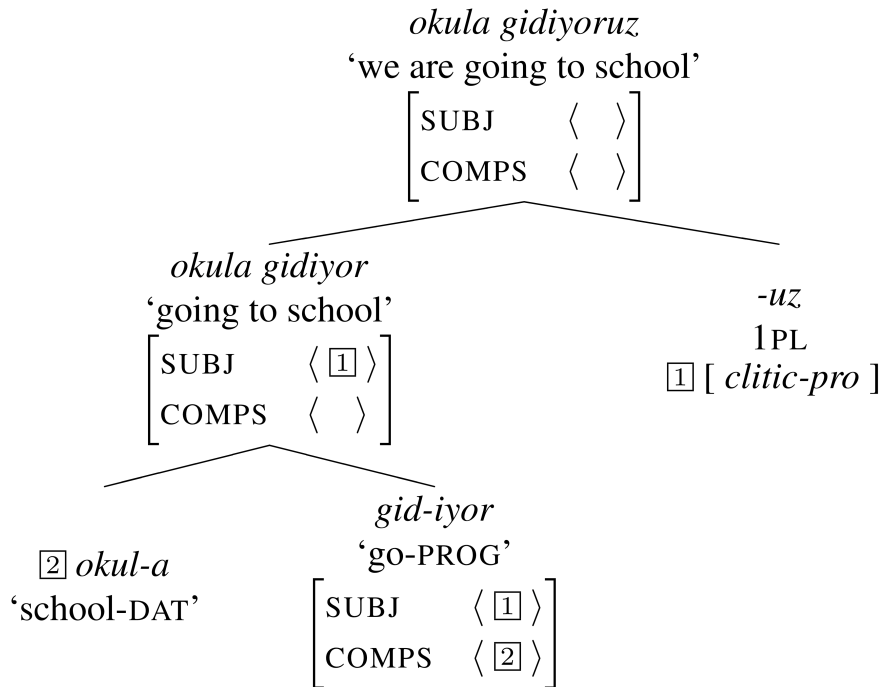
(25)

$$
\begin{bmatrix}
\textit{phrase} \\[4pt]
\text{PHON} \qquad\qquad \begin{bmatrix} \boxed{2} \oplus \boxed{3} \end{bmatrix} \\[4pt]
\text{SUBJ} \qquad\qquad \langle \ \rangle \\[8pt]
\text{HEAD-DGHTR} \qquad \begin{bmatrix} \textit{non-finite-pred} \\ \text{PHON} \qquad \boxed{2} \\ \text{SUBJ} \qquad \langle \boxed{1} \rangle \end{bmatrix} \\[12pt]
\text{NON-HEAD-DGHTRS} \quad \left\langle \boxed{1} \begin{bmatrix} \textit{clitic-pro} \\ \text{PHON} \qquad \boxed{3} \end{bmatrix} \right\rangle
\end{bmatrix}
$$

The constraint in (20) states that the SUBJ value (i.e. subject) of a phrase headed by an element of type *non-finite-pred(icate)* must be of type *clitic-pro(noun)*. In addition, it states that the phonology of the clitic pronoun will appear after the phonology of the non-finite predicate. The type *non-finite-pred* does not appear in the type hierarchy in (19) because it is not a morphological type. Rather, it is a syntactic type which we define as a phrase with a word of type *non-finite* as its head. In the context of this paper, we are mostly concerned with non-finite predicates headed by verbs taking *z*-paradigm endings. However, such predicates would also need to include those headed by predicate nominals and predicate adjectives, as indicated by the data in (3), where *z*-paradigm endings are seen to attach to each of those two types.

An important, and potentially controversial aspect, of the analysis encoded by (20) is that the clitic pronouns in Turkish sentences are taken to be the true syntactic subjects, while the NP's identified as subjects in the traditional 'SOV' descriptions of the Turkish sentences are merely NP's which are coreferential with the subject—we will come back to this issue in section 6.

With (25) in mind, we can now understand how the unification of a predicate and a subject pronominal enclitic works. To illustrate this we give the partially specified structure in (26) where the predicate is a non-finite verb phrase taking a clitic pronoun subject.

(26)   *Analysis of* okula gidiyoruz *'we are going to school'.*

$$
\begin{array}{c}
\textit{okula gidiyoruz} \\
\text{'we are going to school'} \\
\begin{bmatrix}
\text{SUBJ} & \langle \ \rangle \\
\text{COMPS} & \langle \ \rangle
\end{bmatrix}
\end{array}
$$

$$
\begin{array}{c}
\textit{okula gidiyor} \\
\text{'going to school'} \\
\begin{bmatrix}
\text{SUBJ} & \langle \boxed{1} \rangle \\
\text{COMPS} & \langle \ \rangle
\end{bmatrix}
\end{array}
\qquad
\begin{array}{c}
\textit{-uz} \\
1\text{PL} \\
\boxed{1} \, [\ \textit{clitic-pro}\ ]
\end{array}
$$

$$
\boxed{2}\ \textit{okul-a}
\qquad
\begin{array}{c}
\textit{gid-iyor} \\
\text{'go-PROG'} \\
\begin{bmatrix}
\text{SUBJ} & \langle \boxed{1} \rangle \\
\text{COMPS} & \langle \boxed{2} \rangle
\end{bmatrix}
\end{array}
$$

'school-DAT'

Starting from the top node in (26), the left branch of this tree is the sign for the verb phrase 'going to school'. Crucially, its SUBJ list is not empty, which means that to construct a full sentence, the element on the SUBJ list must be matched by an element of the type it specifies for—this must be an element of type *clitic-pro* via the construction given in (25)—which will result in the SUBJ requirement of the verb phrase being fulfilled. Less formally, we can restate this by saying that the verb phrase *okula gidiyor* requires a clitic pronoun subject.

The right branch of (26) is the sign for the first person plural pronominal clitic. As this sign combines with the *okula gidiyor* sign, the resulting sign, the top node of the tree, has both an empty SUBJ list and an empty complement (COMPS) list, which means it is a legitimate sentential sign. Again, less formally, we can restate this idea as: At the level of the sentence all of the verb's arguments must be specified. Finally, we have also drawn out the structure of the verb phrase *okula gidiyor* in (26). However, its internal shape is not critical here.

The analysis exemplified by (26) that we posit for the *z*-paradigm clitics can directly account for the suspended affixation facts seen in (9) and (10). Recall that the *z*-endings are capable of taking wide scope over two or more conjuncts in a coordinated structure. This fact falls out naturally here since we are treating the *z*-paradigm endings as independent signs that are combined with phrases according to canonical syntactic principles. Since coordinated structures
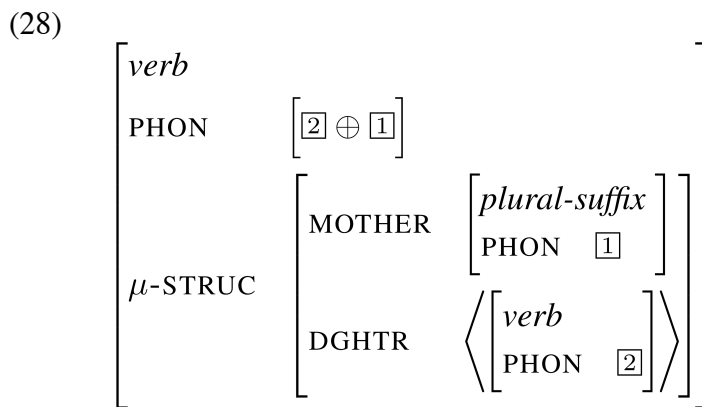
are phrases themselves, the possibility that a *z*-paradigm ending can attach to a coordinated structure is straightforwardly predicted.

4.4.   *The behavior of the plural suffix -lEr*

The analysis we have proposed in the previous sections can be simply extended to account for another important and interesting area of variable affix ordering in Turkish morphology. Good and Yu (2000) observe that the behavior of the plural suffix *-lEr* exhibits an ordering variability similar to what was described for the *k*-paradigm markers in section 1.2. This suffix optionally marks plurality on nouns and verbs with third-person plural subjects. When used on a verb with multiple TMA markers, it can follow any of them, but cannot immediately follow the verb stem, as the data in (27) shows.

(27)   a.   *koş*   *-uyor*   *-du*   *-ysa*   *-lar*
             run   PROG   PAST   COND   **PL**
       b.   *koş*   *-uyor*   *-du*   *-lar*   *-ysa*
             run   PROG   PAST   **PL**   COND
       c.   *koş*   *-uyor*   *-lar*   *-dı*   *-ysa*
             run   PROG   **PL**   PAST   COND
       d.* *koş*   *-lar*   *-uyor*   *-du*   *-ysa*
             run   **PL**   PROG   PAST   COND
             'If they were running…'

Using essentially the same type of construction which accounted for the distribution of verbal TMA suffixes in (20), we can account for the distribution of the plural suffix. The only difference is that, whereas TMA suffixes were formalized as being able to have either elements of type *verb-stem* or *verb* as their daughter, an element of type *plural-suffix* will only be allowed to have a *verb* as its daughter—thus correctly predicting the grammaticality of (27a–c) and the ungrammaticality of (27d).

(28)

$$
\begin{bmatrix}
verb \\
\text{PHON} \quad \left[\boxed{2} \oplus \boxed{1}\right] \\
\mu\text{-STRUC} \quad
\begin{bmatrix}
\text{MOTHER} & \begin{bmatrix} plural\text{-}suffix \\ \text{PHON} \quad \boxed{1} \end{bmatrix} \\
\text{DGHTR} & \left\langle \begin{bmatrix} verb \\ \text{PHON} \quad \boxed{2} \end{bmatrix} \right\rangle
\end{bmatrix}
\end{bmatrix}
$$

We take the fact that the basic pieces of our formalization can be straightforwardly extended to plural ordering variability to lend general support to our analysis.

5. Notes on the morphology-syntax interface in the analysis

Since this analysis is somewhat novel in its combination of the formal devices of Koenig (1999) to describe morphological facts and Sag and Wasow (2000) to describe syntactic facts, a few brief comments should help clarify the sort of interaction between the morphology and syntax we are assuming. Implicit, but not formalized, in the analysis is the commonly-held idea that the syntax and morphology interact at the level of the word. Thus, morphological constructions create syntactic words which can then enter into syntactic constructions. An example of a 'word-creating' construction is the one in (20) which can be instantiated to create verbs from other verbs or from morphological stems which are inaccessible to the syntax.

The two sets of constructions, one morphological and one syntactic, are justified, we feel, by the strikingly different behavior of the two verbal paradigms which can be formalized simply if we assume that morphological constructions govern the use of the *k*-paradigm and syntactic constructions govern the use of the *z*-paradigm.

Our treatment of the Turkish facts presented here markedly departs from more traditional HPSG analyses of the division between morphology and syntax (see, e.g., Sag and Wasow 2000) in that it does not invoke lexical rules to deal with 'morphological' facts and constructions to deal with 'syntactic' facts. Rather, it uses the same basic construction-based architecture to formalize both types of generalizations.

6. Discussion

In our analysis, we have argued that the *z*-paradigm endings should be treated as clitics. A consequence of this is that sentences like (26), whose matrix verbs take these endings, are unambiguously formalized as being OVS. However, Turkish is generally described as being SOV. For example, it is possible in Turkish to have emphatic sentence-initial personal pronouns forming apparent SOV sentences, as in (29a). Furthermore, non-pronominal subjects also generally surface in SOV order, as in (29b).

(29)   a. ***Biz***   *okul-a*   *gid-iyor-uz.*
        1PL   school-DAT   go-PROG-1PL
        'We are going to school'
    b. ***Can***   *okul-a*   *gid-iyor.*
        John   school-DAT   go-PROG
        'John is going to school'

Even when emphatic pronouns are present, like in (29a), pronominal subject marking is required on the verb. Our analysis of Turkish pronominal subject markers combined with the data in (29) suggests that Turkish resembles both an SOV and an OVS language on the surface. The overall implications this conclusion has for the grammar of Turkish are an important matter, and we shall briefly discuss some of them here.

Our formalization implicitly adopts an incorporation-style analysis (cf. Bresnan and Mchombo 1987). That is, we are treating the pronominal clitics as though they are a sort of

incorporated subject pronoun. Given this understanding of the Turkish pronominal morphemes, the problem of 'clitic doubling' naturally presents itself in cases where the overt pronoun surfaces with the pronominal clitics simultaneously, such as the contrastive case given below.

(30) **Ben** *iş-e* *gid-iyor-um, halbuki **o** maç-a gid-iyor-Ø.*
1SG work-DAT go-PROG-1SG however **3SG** match-DAT go-PROG-3SG
'I am going to work, however *he* is going to the game.'

The crucial problem presented by clitic doubling cases is: If the incorporated pronoun is already serving as the subject of the sentence, what role does the overt pronoun have?

If we assume some principle along the lines of the Functional Uniqueness Condition (Bresnan and Mchombo 1987), the overt pronoun cannot be treated as a subject *per se*, since each semantic role can only be associated with one argument (in other words, there cannot be two things functioning as subject at the same time).

One possible way to deal with this issue would be to suggest that the Functional Uniqueness Condition is violable. If this were the case, the problem of clitic-doubling would vanish naturally. One could just stipulate that Turkish allows two arguments satisfying the same semantic role. The claim that the Functional Uniqueness Condition is violable would not be a particularly drastic move in light of the recent advance of Optimality Theory, which strongly advocates the violability of constraints in grammar generally.

Another possible solution to the clitic-doubling problem might be to treat the overt pronoun as a nominal adjunct licensed by certain discourse functions, such as emphasis and contrast. Such claim is not without support and would seem to be the most natural analysis in standard transformational approaches to grammar. A discourse-oriented description of the pronominal system of Turkish is succinctly provided in Erguvanlı-Taylan (1986):

> Turkish marks subject agreement on the verbal element by means of a person suffix; the use of an independent subject pronoun is, then, optional unless the subject has an emphatic or contrastive function, in which case a pronominal form becomes obligatory (Erguvanlı-Taylan 1986: 210).

Free personal pronouns are, otherwise, rarely obligatory. According to Kornfilt (1997), this is due to the fact that the agreement morphemes on predicates make it possible to recover the features of the subject pronoun. Thus, if every time the subject pronoun is omitted, one must 'recover' the features of the subject from the 'agreement morphemes' on the predicates and the head nouns, might it not be simpler to claim that the so-called agreement morphemes are the subject arguments themselves?

However one chooses to deal with this issue, particularly interesting, in the Turkish case, is that, even though *k*-paradigm endings and *z*-paradigm endings systematically differ in many ways, our analysis points in a direction where both are treated as the true subjects of their verbs.

Finally, it would be worthwhile to point out here, that Özturk (2001) has come to some very similar conclusions to the ones we have, with respect to what the 'subject' of the Turkish sentence is, focusing on data unrelated to the data which has been the focus of this paper. Özturk's analysis explicitly treats sentence-initial 'subjects' in Turkish as topics and agreement markers as subjects.

7. Conclusion

In this paper, we have demonstrated that the two subject pronominal paradigms in Turkish display strikingly different morphosyntactic behavior. Providing evidence from morphological selectivity, stress assignment, variable ordering, and suspended affixation, we have illustrated that the *k*-paradigm endings in Turkish can be usefully treated as lexical suffixes, whereas the *z*-paradigm endings are usefully treated as post-lexical clitics. We have also outlined the historical development of the two paradigms, showing that the *z*-paradigm arose from cliticization historically while the *k*-paradigm appears to have developed from the reanalysis of preexisting suffixes—thus demonstrating that our synchronic analysis closely matches the diachronic facts.

Furthermore, we have attempted to sketch out an HPSG account of the data, suggesting that the *k*-paradigm endings can be treated as constrained by a morphological construction in the lexicon while the *z*-paradigm endings can be treated syntactically. The formal analysis is novel within HPSG in using two sets of constructions, a morphological set and syntactic set, in describing the data and, thereby, not making use of lexical rules to handle the morphological facts which were relevant to the syntax.

Our analysis further questions some of the standard assumptions regarding the treatment of subjects in Turkish. Specifically, it runs counter to the idea that Turkish is a straightforward SOV language. Minimally, our analysis indicates that the syntax of word order in Turkish is more complex than is traditionally assumed. More broadly, if our analysis is correct in the details, then it will be an important area of research to establish how precisely one should describe the element of the Turkish sentence typically thought of as being the subject. It will also be interesting to see what implications our analysis of Turkish could have for other 'pro-drop' languages with rich, suffixal agreement morphology which could, in principle, be amenable to an analysis similar to our OVS analysis of Turkish.

Notes

1. More information on the imperative agreement paradigm can be found in Lewis (1967: 137–8) and more on the optative can be found in Lewis (1967: 132–3). The use of these agreement paradigms is much more restricted than the other two, and their properties were not investigated for this paper.
2. Turkish examples given in the text are presented in Turkish orthography. An *ş* represents an [ʃ], a *ç* represents a [tʃ], an *ı* represents an [ɯ], and a *ğ* has no phonetic realization.
3. While this historical account can explain the origin of the past maker and the *k*-paradigm endings generally, it does not explain yet why the first person plural ending of the *k*-paradigm is ends in a *k* as opposed to some sequence like *mIz*. Further research is needed to understand the origin of this part of the paradigm.

References

Adamović, M. 1985. *Konjugationsgeschichte der türkischen Sprache*. Leiden: E. J. Brill.

Bresnan, J. and Mchombo, S. A. 1987. "Topic, pronoun, and agreement in Chichewa". In *Working papers in grammatical theory and discourse structure: Interactions of morphology, syntax, and discourse*, M. Iida, S. Wechsler, and D. Zec (eds.), 1–59. Chicago: University of Chicago Press.

Enç, M. 1986. "Topic switching and pronominal subjects in Turkish". In *Studies in Turkish linguistics*, D. Slobin and K. Zimmer (eds.), 195–208. Philadelphia: John Benjamins Publishing Company.

Erguvanlı-Taylan, E. 1986. "Pronominal versus zero representation of anaphora in Turkish". In *Studies in Turkish linguistics*, D. Slobin and K. Zimmer (eds.), 209–233. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Good, J. and Yu, A. C. 2000. "Affix-placement variation in Turkish". In *Proceedings of the 25th Annual Meeting of the Berkeley Linguistics Society*, J. Good and A. Yu (eds.), 63–74. Berkeley: Berkeley Linguistics Society.

Haiman, J. 1977. "Reinterpretation'. *Language* 53: 312–328.

Halpern, A. (1995). *On the Placement and Morphology of Clitics*. Stanford: CSLI.

Kaşgarli M. 1939–1943. *Divanü Lûgat-it-Türk*. B. Atalay (ed.). I–V. Ankara.

Koenig, J.-P. 1998. *Lexical relations*. Stanford: CSLI.

Lewis, G. L. 1953. *Turkish*. Kent, UK: Hodder and Stoughton.

Lewis, G. L. 1967. *Turkish grammar*. Oxford: Oxford University Press.

Miller, P. H. 1992. "Post-lexical cliticization vs. affixation: Coordination criteria". In *Proceedings of the 28th Annual Meeting of Chicago Linguistic Society*, J. M. Denton et al. (eds.), 382–396. Chicago: Chicago Linguistic Society.

Orgun, O. 1995. "Flat versus branching morphological structures: The case of suspended affixation". In *Proceedings of the 21st Annual Meeting of the Berkeley Linguistics Society*, L. Bilmes, et al. (eds.), 252–261. Berkeley: Berkeley Linguistics Society.

Öztürk, B. 2001. "Turkish as a non-pro-drop language". In Taylan, E. (ed.). 239–259.

Sag I. and Wasow T. 2000. *Syntactic Theory: A Formal Introduction*. Stanford: CSLI.

Sezer, E. 1998. *Yaşayan Türkçe*. Ms. Cambridge, Mass: Harvard.

Sezer, E. 2001. "Finite inflection in Turkish". In Taylan, E. (ed.). 1–45.

Shaw, R. B. 1877. "A grammar of the language of Eastern Turkistán". *Journal of Asiatic Society of Bengal* 46.

Taylan, E (ed.). 2001. *The verb in Turkish*. Benjamins: Amsterdam.

Zwicky, A. and Pullum, G. K. 1983. "Cliticization vs. inflection: English *n't*". *Language* 59: 502–513.

University of Pittsburgh
Department of Linguistics
2816 Cathedral of Learning
Pittsburgh, PA 15260

Jeff Good
jcgood@pitt.edu

Alan C. L. Yu
aclyu@uchicago.edu