Morphosyntax of Two Turkish Subject Pronominal Paradigms

Alan C. Yu and Jeff C. Good

University of California, Berkeley

0. Introduction

This paper begins with a presentation of a split in the morphosyntactic behavior of two suffixing subject pronominal paradigms in Turkish in section 1. In section 2, we argue that this split is a result of one paradigm consisting of postlexical clitics while the other is composed of lexical suffixes. To better appreciate the distinct morphosyntactic behavior of these two paradigms, we will present a brief overview of the historical development for such bipartite behavior in section 3. We will then present an account of it in a lexicalist framework, namely Head-driven Phrase Structure Grammar (HPSG), in section 4. A brief conclusion appears in section 5.

1. Pronominal Subject Suffixes in Turkish

Turkish is a Turkic language, spoken mainly in Turkey and its neighboring countries. It should be noted that data used in this study is based on judgements of speakers of the Istanbul dialects of Turkish. The forms used for pronominal subject paradigms can vary greatly across dialects.

* A preliminary version of this paper was presented at the Workshop on the Perspectives on Clitic and Agreement Affix Combinations, University of Illinois, Urbana-Champaign, July 28, 1999. We would like to thank the participants at the workshop for their comments. We would also like to thank Andreas Kaithol, Larry Hyman, and Sharou Inkelas for their comments on drafts of this paper.

© 2000 by Alan C. Yu and Jeff C. Good
NELS 30
1.1. The Basics

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 since they bear no relevance to the present discussion. The two paradigms that we will focus on are given in (1).\(^1\)

<table>
<thead>
<tr>
<th></th>
<th>(k)-paradigm</th>
<th>(z)-paradigm</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1^\text{st})</td>
<td>-m</td>
<td>-(y)Im</td>
</tr>
<tr>
<td>(2^\text{nd})</td>
<td>-n</td>
<td>-sIn</td>
</tr>
<tr>
<td>(3^\text{rd})</td>
<td>-Ø</td>
<td>-Ø</td>
</tr>
</tbody>
</table>

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

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>a. dön-dü-m</td>
<td>b. dön-se-m</td>
</tr>
<tr>
<td></td>
<td>dön-dü-n</td>
<td>dön-se-n</td>
</tr>
<tr>
<td></td>
<td>dön-dü-k</td>
<td>dön-se-k</td>
</tr>
<tr>
<td></td>
<td>dön-dü-nüz</td>
<td>dön-se-nüz</td>
</tr>
<tr>
<td></td>
<td>turn-PAST-PSN</td>
<td>turn-COND-PSN</td>
</tr>
</tbody>
</table>

The other paradigm (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.)

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>gid-iyor-üz</td>
<td>‘we are going’</td>
</tr>
<tr>
<td>b.</td>
<td>adam-iz</td>
<td>‘we are men’</td>
</tr>
<tr>
<td>c.</td>
<td>iyi-yiz</td>
<td>‘we are fine’</td>
</tr>
<tr>
<td>d.</td>
<td>*git-ti-yiz</td>
<td>‘we went’</td>
</tr>
</tbody>
</table>

(3a-c) show the types of predicates the \(z\)-paradigm can attach to. The corresponding ungrammatical forms with the \(k\)-subject pronominal suffixes are shown at the end of each example. (3d) demonstrates that \(z\)-paradigm subject pronominal markers

---

\(^1\) Turkish examples given in the text are presented in Turkish orthography. An ğ represents an [i], a ğ represents a [ç], an ı represents a [i], and a ş represents a marker of vowel length or hiatus.
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 mentioned so far are the most straightforward differences between the two paradigms. However, there are actually more idiosyncrasies that demonstrate a major morphosyntactic distinction between them. We shall examine these facts one by one.

1.2. Suffix Order Variation

In all the examples above, subject pronominal markers were the final suffix 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 subject pronominal suffixes 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. That is, the k-paradigm endings can appear predicate-finally or between the two TMA markers. Such variability is not possible for the z-paradigm suffixes.

(4) a. gör - dü - yse - m
    see-PAST-COND-1SG
    'if I saw ...'
    gör - dü - m - se
    see-PAST-1SG - COND

    b. gör - dü - yse - n
    see-PAST-COND-2SG
    gör - dü - n - se
    see-PAST-2SG - COND

    c. gör - dü - yse - k
    see-PAST-COND-1PL
    gör - dü - k - se
    see-PAST-1PL - COND

    d. gör - dü - yse - núz
    see-PAST-COND-2PL
    gör - dü - núz - se
    see-PAST-2PL - COND

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 license conflicting pronominal paradigms. In this case, we have a combination of the progressive marker, which licenses only the z-paradigm endings and the past marker, which licenses only the k-paradigm endings.
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 license conflicting pronominal paradigms. However, as the data in (6) illustrate, even when both of the TMA markers are $z$-paradigm licensors, $z$-paradigm endings must still surface at the end of the predicate.

This inability of the $z$-paradigm 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 host selectivity. These two paradigms systematically differ across a range of linguistic parameters which demands a much more principled explanation than merely attributing such morphological idiosyncrasies to chance. In the remainder of this paper, we will explicate the reasons for this bipartite behavior on both synchronic and historical grounds. We will also present a formal account that attempts to capture the observed generalizations.

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

The difference between the $k$- and the $z$-paradigm, as we shall argue, is a matter of their formal status in the lexicon. That is, we assert that the $k$-paradigm endings are lexical suffixes but the $z$-paradigm endings are postlexical clitics.
To verify that the z-paradigm endings are indeed post-lexical clitics or phrasal affixes (cf. Anderson 1992), we rely on the diagnostic conditions that were provided in Zwicky and Pullum (1983). The 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. k-endings only follow two verbal suffixes while z-endings follow all other verbal suffixes as well as non-verbal predicates. The variable ordering of k-paradigm suffixes 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, treats verbs with k-endings as whole constituents whereas it does not treat the combination of verb + z-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 do not usually receive lexical stress. As the data in (8) shows, this generalization is also borne out here. Default Turkish stress is word-final. Forms with k-paradigm endings can be stressed when they are word-final whereas z-paradigm endings can never be stressed. Thus, word-final k-paradigm endings behave as though they are truly suffixed to the word, as opposed to the forms with the z-paradigm endings where stress always lands on the preceding syllable.
So, looking at the data in light of the criteria set forth by Zwicky and Pullum, helps us justify 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 having wide scope over a conjunction of hosts. This, as we shall show, is also allowed for the z-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, 1996). It is a construction in Turkish where suffixes are optionally omitted from all conjuncts in a coordinated structure except for the last one. The suffixes 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 morpho-syntactic distribution of the subject pronominal endings.

\[(a) \quad \text{[genç ve büyüğ]-üm}\]
\[\text{[young and big]-1SG}\]

\[(b) \quad \text{[genç]-im ve [büyüğ]-üm}\]
\[\text{[young]-1SG and [big]-1SG}\]

'I am young and big'

\[(c) \quad \text{[hastane-ye gid-iyor, o-nu gör-iyor]-sunuz}\]
\[\text{[hospital-DAT go-PROG 3SG-ACC see- PROG]-2PL}\]

\[(d) \quad \text{[hastane-ye gid-iyor]-sunuz, [o-nu gör-iyor]-sunuz}\]
\[\text{[hospital-DAT go- PROG]-2PL 3SG-ACC see- PROG]-2PL}\]

'You all are going to the hospital and seeing him/her.'

(9a) illustrates that the first person singular suffix -(y)l in 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 suffix -(y)lz in (9c) has scope over both go
Morphosyntax of Two Turkish Subject Pronominal Paradigms

and see. No personal agreement suffix appears on the first conjunct, go, since the personal agreement is realized on the second conjunct, see. The unsuspended version of (9c) is given in (9d). The application of this suspended affixation construction, however, is not totally unrestricted. Note that both of the personal agreement suffixes in (9a, c) are personal suffixes of the z-paradigm. k-suffixes are prohibited from participating in this suspended affixation construction.

(10) a. [[ev-e gel-ir] [sana yardm 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 yardm et-ii]]-k
    [[home-DAT come-PAST] [you.DAT help do-PAST]]-1PL

c. [[ev-e gel-di-k] [sana yardm 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) as the subject pronominal ending required is from the z-paradigm. (10b), on the contrary, is ill-formed as the subject pronominal suffix suspended is from the k-paradigm. The k-paradigm suffix is required there since the simple past tense suffix -(y)DI only takes suffixes from this paradigm. The well-formed equivalent of (10b) is given in (10c). The fact that suspended affixation is impossible for the k-paradigm suffixes suggests there is strong lexical affinity of the k-paradigm suffixes to the simple past suffix -(y)DI. These same basic facts are true for the conditional suffix -(y)sE.

3. Historical Perspective

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

Old Turkic did not have suffixing subject agreement markers. Sentences with pronominal subjects were formed by putting a pronoun at the end of the predicate.
The first cliticized forms of the predicate final pronouns appeared in 13th century texts. A typical paradigm during that period is given in (12).

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

Although the historical development of the $z$-paradigm is rather transparent and simple, the historical origin of the $k$-paradigm is still rather unclear. Shaw (1877) pointed out that the Old Turkic preterite was formed via the possessive construction as in (13).

The morpheme $-d$ in (13) was serving as a nominalizing suffix while the $-um$ suffix was the first person singular possessive suffix. According to Adamovic (1984), a reference in Kasgarli (1939:60-63) mentioned that the preterite was in mutually exchangeable relations with the periphrase formed by the deverbal nominal suffix $-dug$/ $du$/$dik$ in Oghus, Suwar and probably Kipchak in the 11th century. This is exemplified in (14).
This historical account mostly predicts the origin of the past maker and the k-paradigm endings. However, it does not explain why the first person plural ending of the k-paradigm is -dik and not -dumuz. We have yet to find a satisfactory historical account of this change. However, we can still appreciate the morphosyntactic status of the predicate formed by the verb, the past tense suffix and the k-paradigm endings. That is to say, the k-paradigm endings were never 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 which have been reduced, in several stages. This opposing historical development offers an account of how the synchronic split of the k-paradigm endings as suffixes and the z-paradigm endings as clitics developed.

Assuming the historical development of the k-paradigm in the preterite is valid, the application of the k-paradigm to the conditional predicate appears to be a case of paradigm levelling at a later stage in Turkic. This levelling may have been motivated by the fact that these are the only two verbal suffixes that end in vowels.

4. Analysis

For the remainder of this paper, we will briefly sketch out a formal account for the data. However, before we dive into the discussion, we shall briefly review the generalizations that we should capture in formal terms. They are summarized in (15).

(15) 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 suffix 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.

4.1. Morphological vs. Phrasal Realization of the Subject Paradigms

Our analysis will be couched in the framework of Head-driven Phrase Structure Grammar (Pollard & Sag 1994), which is a unificational constraint-based grammar. In order to account for the morpho-syntactic distributional dichotomy of the \( k \)- and \( z \)-paradigms, we propose first to account for the realization of the \( k \)-paradigm by means of a morphological schema (cf. Riehemann 1998; Koenig 1999). That is, the affixing of the subject pronominal for predicates that take the \( k \)-paradigm must take place in the lexicon. On the other hand, we will capture the clitic nature of the \( z \)-paradigm by making use of syntactic constraints.

The verbs that only take the \( k \)-paradigm are subject to the constraint below.

\[
(16) \quad \begin{array}{|l|c|c|c|}
\hline
\text{MORPH} & \text{FORM} & F([0], ...) \\
\text{I-FORM} & [0] & \\
\text{SUBJ-K-VB} & \text{SYNSEM} & \text{HEAD} \\
\text{VAL} & \text{COMPS} & [2]\text{list(canon-ss)} \\
\text{ARG-ST} & <[1]\text{mark}\@[2] & \\
\hline
\end{array}
\]

The schema in (16) says that the subject, [1], is of type *affix*. This information is given in the specification of the first member of the ARG-ST list. This constraint also says that the verb can have a list of complements that are of type canonical syntax-semantics (cannon-ss)—that is, nonaffixal. This implies that all complements of the verb must be full lexical items or phrases, which is the case in Turkish since it does not mark for objects on the verb. Notice that the HEAD value is specified as being of the type \( k \)-verb. \( k \)-verb is a general type that encompasses subtypes such as the \( past \)-\( vb \) and the \( conditional \)-\( vb \). These subtypes are themselves morphologically complex as they are composites of a verb stem plus the TMA inflectional marker.

Now, as for the phonological realization of the individual \( k \)-paradigm endings, we follow Müller and Sag 1997 by assuming that it is determined by a function \( F \), which requires that the FORM value be related to the (inflected)-FORM value via the appropriate suffixation. Unfortunately, we do not have space to go over in detail how such function would work. The central idea behind using such a function is to encode that \( k \)-paradigm endings are added as the result of some morphological function which
attaches the endings in the lexicon. This is in opposition to z-paradigm endings which are added in the syntax by a phrase structure rule.

Note that this formulation of the k-suffixation also prohibits the possibility of the k-paradigm from participating in suspended affixation since conjunction is a syntactic operation while the k-suffixation takes place in the lexicon.

The variation in the placement of the k-endings when the last two TMA markers are conditional and simple-past is accounted for by our analysis straightforwardly. Since k-suffixation happens in the lexicon, there is nothing to prohibit words with a k-ending appearing between two k-TMA markers, since the constraint in (16) will be satisfied either way. This is schematized in (17).

(17) Formation of <üyüdü, yse, k> 'If we slept...'

a. [üyüdü-yse]-k | VAL | SUBJ < | ARG-ST <[1]_kΦ [2]>


b. [üyüdü-k]-sc | VAL | SUBJ < | ARG-ST <[1]_kΦ [2]>


Finally, regarding the treatment for the z-paradigm endings, recall that we would like to have a way to account for the z-endings' realizations phrasally. To do this, we propose to treat the z-enclitics as independent signs that are combined with verbal or nominal predicates through mechanisms in the syntax. We posit that the signs that characterize the z-enclitics should be like the example in (18). This pronominal enclitic sign has three crucial specifications: The phonology that relates the phonological instantiation of the z-ending to its host and the person and number features of the particular z-ending.
This sign can unify with any predicate that is looking for a clitic pronominal subject. In order to demonstrate how the unification works, it is necessary to formalize how a given predicate specifies what subject value it should take. We posit that all words that can serve as predicates can take on the form shown in (20).

(20) \[ \text{pred-wd} \Rightarrow \begin{array}{c} \text{SUBJ} <\text{clitic-pro}> \\ \text{ARG-ST} <\text{NP}>@\text{list}(<\text{canon-ss}> ) \end{array} \]

This constraint states that the SUBJ value of a predicate-word is of type clitic-pro(noun). This is co-referent with the first NP of the ARG-ST—the first member of the ARG-ST list is always the subject in HPSG. With (20) in mind, we can now understand how the unification between a predicate and a subject pronominal enclitic works. To illustrate this we give the partially specified structure in (21).

(21) The sign for the sentence okula gidiyoruz 'we are going to school'.

```
gidiyoruz 'we are going to school'
  SUBJ < >
  COMPS < >
  gidiyor 'going to school'
    SUBJ < [1]clitic-pro >
    COMPS < >

    SUBJ < [1] >
    COMPS < [2]>[enclitic]
```

Starting from the top node, 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 a phrase of the
type it specifies for, here clitic-pro, which will result in the SUBJ requirement of the verb phrase being cancelled. The right branch is the sign for the first person plural pronominal clitic. As the two signs combine, the resulting sign has both an empty SUBJ and an empty COMPS lists, which means this is a legitimate sentential sign. We have drawn out the structure of the verb phrase in this analysis. However, its internal shape is not critical here.

The analysis that we posit for the z-paradigm clitics has the benefit of accounting for the suspended affixation facts. 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-endings as independent signs that are combined with phrases according to canonical syntactic principles. Now, since coordinated structures are phrases themselves, the possibility of a z-paradigm subject pronominal to attach onto a coordinated structure is predicted.

A consequence of this analysis is that sentences like (21), which take z-paradigm endings, are formalized as being OVS. However, Turkish is generally described as being SOV. It is possible in Turkish to have emphatic sentence-initial personal pronouns forming SOV-like sentences, as in (22a). Furthermore, non-pronominal subjects also generally surface in SOV order, as in (22b).

(22) a. biz okul-a gid-iyor-uz
    1PL school-DAT go-PROG-1PL
    \textit{\textquoteleft we are going to school\textquoteright}

    b. Can okul-a gid-iyor
    John school-DAT go-PROG
    \textit{\textquoteleft John is going to school\textquoteright}

Importantly, even when emphatic pronouns are present, like in (22a), pronominal subject marking is required on the verb. Our analysis of Turkish pronominal subject markers combined with the data in (22) suggests that Turkish resembles both an SOV and an OVS language on the surface. Unfortunately, we do not have the space here to properly discuss the overall implications this conclusion has for the grammar of Turkish.

5. 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 suffix ordering and suspended affixation, we have illustrated that the k-paradigm endings in Turkish should be treated as lexical suffixes, whereas the z-paradigm endings should be treated as clitics. We have
also outlined the historical development of the two paradigms, showing that the z-paradigm arose from cliticization historically—thus demonstrating that our synchronic analysis closely matches the diachronic facts.

Finally, 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 schema in the lexicon while the z-paradigm endings can be treated syntactically.

References

Morphosyntax of Two Turkish Subject Pronominal Paradigms

Department of Linguistics
1203 Dwinelle Hall
University of California
Berkeley, CA 94720-2650

aclyu@socrates.berkeley.edu and jgood@socrates.berkeley.edu