The theoretical background

The theory of lexical phonology as worked out in Kiparsky (1982), Mohanan (1982), Pesetsky (1979) and others forces a fundamental division of rules. Rules that require lexical information for their application apply in the lexicon. Phonological rules that do not depend on morphological categories may apply either lexically or postlexically. Kiparsky (1983) summarizes the main difference between lexical rules and postlexical rules in the form of a table given below:

<table>
<thead>
<tr>
<th>Lexical rules</th>
<th>Post-lexical rules</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word bounded</td>
<td>not word bounded</td>
</tr>
<tr>
<td>Access to word-internal structure at the same level</td>
<td>access to phrase structure only</td>
</tr>
</tbody>
</table>
In this paper I attempt to develop a model of lexical phonology for Modern Standard Turkish (henceforth MST). I will show that MST requires a lexical model that consists of three levels within both phonology and morphology. It will be shown that the MST stress rules of Initial Stress Assignment (ISA) and Final Stress Assignment (FSA) apply at levels 1 and 2, respectively. I shall examine the phonology of two types of compounds in MST: subcompounds and cocompounds. These compounding processes are assigned to distinct levels on the grounds that subcompounds, but not cocompounds, undergo the rules of Vowel Deletion and $\bar{V}$-Deletion. The rule of FSA, followed by the Stress Reduction Convention (SRC), will account for stress assigned to compounds as well. I will argue that the model of level-ordering developed in this paper is both explanatorily and descriptively more adequate than the arbitrarily stipulated boundary symbols.

2.0 The level structure

I assume MST to have three levels of word formation: (1) derivational and reduplication, (2) derivational, inflectional and subcompounding, and (3) inflectional, derivational and cocompounding. These levels are ordered as follows:
Level 1: derivations and reduplication
Level 2: derivations, inflections and subcompounding
Level 3: inflections, derivations and cocompounding

The set of morphological rules of affixation whose domain of application is specified in terms of the levels given above exhibit differences in stress assignment. Therefore, the reason for making this level distinction within the morphology is that the attachment of level 1 morphological categories triggers word-initial stress placement, whereas those of level 2 require stress to be assigned word-finally. (The rule of Final Stress Assignment is given in 2.2, and the rule of Initial Stress Assignment in 2.3.) Moreover, subcompounding processes assigned to level 2 undergo the rules of Vowel Deletion and h-deletion, but the co-compounding processes of level 3 do not. (These two rules are given in 3.0.)

Of the three levels, the first level comprises the directional adverb-forming suffix -rE, the instrumental case suffix -in, a restricted number of zero-derived adverbs and a set of intensive adjectives created through the only reduplicated prefix. Also incorporated in the morphology of level 1 are other adjectives which have an intensified meaning formed via a morphological rule attaching the diminutive suffixes -cik, -icik, -cak and -acik to the root. Vocative and interjection processes also belong here.

To the second level of morphology, I assign derivational suffixes like -lik, which derives abstract nouns from adjectives, and -ci, which is added to nouns to derive other nominals denoting a person professionally or habitually associated with the indicated property. This level also includes other derivational suffixes like -siz, which is equivalent to the English suffix -less or 'without', and the relative -ki 'which'. Moreover, level 2 incorporates the following list of inflectional suffixes: -ler 'plural', the possessive suffixes (-im 'my', -in 'your (sing.)', -sin 'her/his/its', -inlz 'your (plur.)', -lerin 'their'), the locative -DE, the genitive -IN, the accusative -li, and the ablative -DEN. The third level contains the following enclitic suffixes: the personal endings (-im 'I', -sin 'you', -(dir) he/she, -iz 'we', -snlz 'you (plur.)', -dir/ler 'they'), the negative particle -ME, the interrogative particle -ni, the adverbial suffix -CE, the adverbial suffix -ken 'while', the adverbial -leyin 'by or at', the nominal -IL, which derives adjectives from nouns, the progressive aspect -yor, the narrative past indicating reported speech, and the suffix -da 'too'.

The morphemic distinction given above is made in terms of levels which characterize the domain of morphological, semantic and phonological rules. Every undervived lexical item has to enter the system at level 1. Given the feeding nature of the levels, an output from level 1, for example, may undergo the phonological and morphological rules of level 2. In what follows, I will show that the level distinction made within the morphology is necessitated by the application of the stress assignment rules of MST. I will also show how the proposed level-ordered model for the morphology of Turkish ensures the proper order of attachment of affixes.

2.1 Stress Assignment

Stress rules of MST are not sensitive to differences in syllable weight. Generally, words in MST are characterized by the fact that the final syllable is singled out as the bearer of main stress. In 2.2, I will show that, for each successive affix attached at level 2, the stress becomes correspondingly further removed from the root. Thus the stress falls on the final syllable. This predictable and pervasive stress placement can be represented by an unbounded tree "where the head governs the entire sequence of units... on its left" (Halle & Clements 1983:17).

Stress assignment in MST is characterized as follows:
Over the syllables of the word, construct a right-headed unbounded tree:

```
S  ~  S  ~  S
W  W  W  W  S
V  V  V  V  V
```

2.2 Level Phonology and Morphology

A striking fact about Turkish is that, when the root is extended through a level 2 process of agglutination, the ultimate stress shifts to every newly added suffix. This pervasive process of final stress assignment is exemplified in the examples in (1).²

(1) deve 'camel'
develer 'camels'
develerim 'my camels'
develerimiz 'our camels'
develerimizde 'on our camels'

Final stress is assigned by the rule in (2).

(2) Final Stress Assignment (FSA).

Assign word final stress at the end of level 2.

The domain of FSA is the termination of level 2. This has the effect of making the rule apply in a non-cyclic mode.

Adopting the terminology of SPE, I regard level 2 suffixes as being stress non-neutral. Suffixes attached at level 2 trigger the level 2 phonological rule in (2), which assigns final stress. The derivation in (3) illustrates the application of Final Stress Assignment. For ease of presentation of lexical derivations, rules of vowel harmony, final devoicing and voicing assimilation are applied throughout the derivations, even though they may actually reside in the post-lexical component.³

(3) [çocuk] underived lexical item
    [[çocuk]lar] -ar suffixation
    Level 2
    [çocuklar+miz[ım]z] BEC

'of our children'

Another option is to allow FSA to reapply after every successive addition of a suffix. This process sets up the rule as a cyclic one.⁴ A cyclic application of FSA would create a cumbersome and a needlessly redundant process, and hence I have opted for a single level final FSA.

2.3 Level 1 Phonology and Morphology

The affixes common to level 1 have already been stated in 2.0. The words resulting from the morphological and phonological operations at level 1 are a small and select group of adverbs, intensified adjectives, interjections and vocatives. When level 1 affixes are added, initial word stress is assigned by the rule in (4).

(4) Initial Stress Assignment Rule (ISA):

Assign word initial stress in a derived environment (after an affix has been added).
Notice that the affixes at level 1 are either phonetically realized suffixes, zero suffixes or a prefix which is created by a partial reduplication of the initial syllable. Despite the varied types of affixes, the phonological rule of ISA need not contain any morphological information other than the imposition of a derived environment on the rule application. The domain of ISA is defined to be level 1, which allows the rule to apply only to forms derived at level 1 such as bu-ra 'this place', su-ra 'that place', o-ra 'that place', ne-re 'what place'.

(5)  

```
Level 1.  [[bu]ra] --- ISA  
         [[[bu]ra] --- ISA  
         [[bʊra] --- BEC  
```

Affixes at level 1 have the property of triggering the level 1 phonological rule which assigns stress to the first syllable of the derived item. Affixes at level 2 trigger the level 2 phonological rule which assigns final stress. However, when a level 1 derived item acquires a level 2 suffix the stress assigned at level 1 remains on the initial syllable. This phenomenon is illustrated in (6).

(6) nérade 'where'  gradan 'from there/that place'

The items in (6), which have gained word initial stress at level 1, now will have another stress resulting from the application of FSA at the end of level 2. At the end of a level, a lexical item cannot maintain two main stresses. A stress reduction convention will apply to reduce the final stress in words which have gained two main stresses. The Stress Reduction Convention (SRC) is stated as in (7).

(7) Stress Reduction Convention (SRC):
```
In a string which has been assigned two stresses, final stress is reduced. (Domain: level 2)
```

When there are two competing stress rules assigning stress at both levels 1 and 2, respectively, the initial stress survives the SRC, which is a manifestation of a universal constraint that forbids more than one primary stress in a given string. The application of the rules in their domains is shown in (8).

(8)  

```
Level 1.  [[ne]re] [o]re] --- ISA  
         [[[ne]re] [o]re] --- ISA  
         [nére] [ore] BEC  
```

LEXICAL PHONOLOGY AND MORPHOLOGY OF MODERN STANDARD TURKISH
79
With respect to stress assignment, level 1 affixes override level 2 suffixes. It is this unique nature of level 1 affixes which defines them as the closest of all affixes to the root.

Zimmer (1970) investigates the non-final stress pattern of intensive adjectives and vocative forms in MST. Initial stress assignment to these forms, he states, is an exception to the usual word final stress assignment rule of the language. In his treatment, the cases which exhibit word initial stress are considered "syntactically or semantically marked categories." He continues:

"It seems to me, in a rather vague and intuitive way, that both the intensive adjectives and the vocative forms of nouns are somehow emphatic...it can at least be said that the intensive adjectives constitute the marked category of the opposition intensive vs. non-intensive, and that vocatives constitute a marked category as opposed to the syntactic cases" (p.162).

Zimmer’s recourse to the markedness notion describes the facts but does not explain them. In this paper, I account for initial stress in MST words by the ordering of levels in a structured lexicon.

2.4 Level 3 Phonology and Morphology.

The class of suffixes which attach at level 3 have been listed in 2.0. Also, the domains of word stress rules are assumed to be levels 1 and 2. This has the effect of making level 2 derivations which undergo morphological rules of suffixation at level 3 maintain the stress assigned to them at level 2. This is illustrated by the derivation of the items in (10):

(9) Level 3

a. [okú]  b. [gelf]  U.R.
   [[[okú]ma][s+n]  [[[gelf][yor][me]]  Suffixation
      ---  ---  SRC
      [okúmas+n]  [gelf[yor][me]]  BEC
      'let him not read'  'is he coming'

c. [gel]  d. [guzel]  UR
   [[[gel][me]]  [[[guzel][ce]]  Suffixation
      ---  ---  SRC
      [gel][me]  [guzel][ce]  BEC
      'do not come'

e. [gece]  UR
   [[[gece][leyin]]  Suffixation
      ---  SRC
      [gece][leyin]  BEC
      'by night'
The attachment of affixes subject to varied stress rules (two rules of Turkish and the Stress Reduction Convention) is correctly predicted by the level ordering of morphology and phonology. (For further examples of all possible derivations, see the Appendix). Notice that, although there are only two stress rules and the SRC, there are three levels. The stress rule at level 1 is unique to those level 1 affixes. Should an underived form not acquire a level 1 affix and thereby escape level 1 stress, it must obtain level 2 final stress. Level 2 derivations which undergo morphological rules at level 3 maintain their stress assigned at level 2. This is explained by the lack of a stress assignment rule at level 3.

2.5 Stress in foreign words.

Words from foreign languages have inherent stress, and that stress comes from the original language. Consider the examples in (11), which are taken from Dobrovolsky (1976,87).

(11) rádyo 'radio'
kökteyl 'cocktail'
fútboł 'football'
kánada 'Canada'
lókánta 'restaurant'
iskónto 'discount'
ingilêtre 'England'
kerávyé 'caraway'

It is interesting to note that, when the borrowed word has final stress, it is reanalysed and consequently the stress migrates progressively rightward whenever a level 2 suffix is attached, as in (12):

2) [otobús] 'bus'
[(otobús)lér] 'buses'
[(otobús)lér]də 'on the buses'

3.0 Compounds.

In this section, I will consider two types of compounds in MST: endocentric and exocentric compounds. Endocentric compounds have a head which gives the basic meaning of the compound as a whole. Exocentric compounds are not headed in the same way. The head of an endocentric compound provides the thrust of the combined meaning of the two words comprising the endocentric compound. For example, post-ane 'post office', which consists of posta 'mail' and hane 'house', means a place where only mail is handled. In an exocentric compound, the meanings of both items are engaged equally in producing the total, non-compositional meaning. The meaning often goes beyond the sum of the two parts. For example, kahvehane 'coffee house', which is composed of kahve 'coffee' and hane 'house', is not a place where only coffee is served but rather is a restaurant-type establishment.

Adopting the terminology of Mohanan (1982), I will call the two types of compounds subcompounds and cocompounds, respectively. The two compounding processes in MST should be assigned to distinct levels. One reason for distinguishing subcompounds from cocompounds is that the former, but not the latter, does not permit a VV sequence to surface. This is illustrated in (13a):
Vowel deletion takes place in (13a) but not in (13b). The rule of Vowel Deletion is stated in (14) as a mirror image rule:

(14) Vowel Deletion:

\[ V \rightarrow \emptyset \]

The processes of Subcompounding and Cocompounding both set up the phonological environment in which the rule of Vowel Deletion operates. Vowel Deletion, therefore, would be expected to apply in both subcompounds and cocompounds. However, only subcompounds undergo Vowel Deletion. The rule would be expected to apply in cocompounds as well, yielding a wrong output, unless the compounding processes are assigned to distinct levels. I propose that Subcompounding and Vowel Deletion have level 2 as their domain of application, and that level 3 defines the domain of Cocompounding. Therefore, the output of the cocompounding formation rule is exempted from the phonological rule of Vowel Deletion, as desired.

MST has another rule applying in subcompounds but not in cocompounds. The rule deletes an /h/ between two identical vowels (see the examples in (13a)). This rule, which I shall refer to as h-Deletion, is given in (15):

(15) h-Deletion: 

\[ h \rightarrow \emptyset \]

Cocompounds do not undergo (15) as the examples in (13b) show. This fact is accounted for by assuming that rule (15) applies at the subcompounding level, which is level 2. The rule of h-Deletion further supports the need for the level division for compounds. Cocompounds are more loosely combined at level 3 than those combined at level 2. Level 2 subcompounds are more...
tightly bound, a fact which has the consequence of making them vulnerable to phonological rules. The level division within compounds correlates with a difference in their productivity. Level 3 compounds tend to be more productive. This is due to the absence of a phonological restriction on their formation, that is, they are not subject to the constraint imposed by the application of the phonological rules of Vowel Deletion and/or h-Deletion.

The lexicon of MST can accordingly be schematized as follows:

3.1 Compound Stress.

In a compound, only the last syllable of the first member is stressed.

(16) bağ 'head' bakan 'minister' başbakan 'prime minister'
toplulü 'knobbed' ışın 'needle' toplulüışın 'pin'
vakit 'time' kaybetmek 'waste' vakıt[tabetmek 'waste time'
meslek 'profession' das 'mate' meslek[as 'colleague'
kara 'black' yel 'wind' kara[el 'northwester'

The rule of FSA (2) and the SRC (7) provide an account for how stress can be assigned to a compound. The application of Final Stress Assignment at level 2 stresses the final syllable of each member of a compound. Then the SRC applies after the morphological process of compounding, destressing the final stress, as shown in (17):

(17) [ba~] [bakan] underived lexical item
[ba~] [bakan] FSA Level 2 [ba~bakan] Subcompounding [ba~bakan] SRC

Similarly, cocompounds are derived at level 3. The two elements acquire final stress at level 2; they are combined at level 3 and thus the cocompound undergoes the SRC at level 3.

4.0 Boundary Solution and Level Ordering.

In the remainder of this paper, I will show the inadequacy of the use of boundary symbols in making morphological information available to phonological rules. The function of the distinct boundary symbols is to encode morphological information in
will not apply if its structural description does not contain a boundary which is present in a phonological string $K$.

If a boundary solution were assumed, then one would presumably posit a primary boundary (+) corresponding to level 1, a secondary boundary (#), to level 2, and a double boundary (##), to level 3. The boundary would activate the rule of Initial Stress Assignment. The # boundary would allow the application of Final Stress Assignment, Vowel Deletion and $n$-Deletion, but the ## boundary would block the application of these rules.

The imposition of the above-mentioned boundary symbols is arbitrary and inherently unmotivated. Moreover, it would be difficult to include the boundary symbols since level 1 contains both prefixes and suffixes which trigger word initial stress. All other stress is either word or stem final.

In lexical phonology, the triggering effect of boundaries is obtained by defining the domain of a phonological rule $X$ as a level and replacing the boundaries with brackets. The blocking effect of boundaries is achieved by defining the application domain of a phonological rule $Y$ to be a level which is ordered prior to the morphological process that blocks rule $Y$. By grouping together morphological rules with the associated phonological rules, the phonological rules apply naturally to the output of the morphological rules. In this way phonological rules access morphological information directly. Moreover, level-ordering ensures the proper attachment of affixes. It predicts that suffixes of higher levels do not attach to stems which have acquired suffixes from earlier levels. For example, (18a) is well-formed but (18b) is not.

   'right here'

In the lexical phonology model discussed in this paper, no ad hoc stipulation is required on the applicability or non-applicability of rules.

APPENDIX

1. Level 1 derivations (attachment of affixes at level 1):
   a) $-CIG$ is a diminutive suffix. Recall that rules of vowel harmony and voicing assimilation are included in the derivations.

   
   $kisa$ $[u]fa$ $[ev]$ underived lexical items
   $[kfas]cik$ $[ufak]$ $[ev]cik$ ISA
   $kfas[ci]k$ $[ufak]$ $[ev]cik$ BEC
   'very short' 'tiny' 'little house'

   b) $-in$ instrumental case suffix:
   $[yaz]$ $[kis]$ $[ansiz]$ $[guz]$ underived lexical items
   $[yaz]in$ $[kisin]$ $[ansiz]in$ $[guz]in$ -In suffixation
   $[yazin]$ $[kisin]$ $[ansizin]$ $[guz]in$ ISA
   $yazin$ $[kisin]$ $[ansizin]$ $[guzin]$ BEC
   'in summer' 'in winter' 'suddenly' 'in autumn'

   c) $-@$ suffix, deriving adverbs from nouns and adjectives. Consider the nouns and adjectives in (i) as contrasted with their derived adverbs in (iii):

   (i) $nfhayet$ 'end' $nfhaye$-@ 'finally'
   $cokluk$ 'multitude' $cokluk$-@ 'often'
   $artik$ 'residue' $artik$-@ 'thereafter'
   $yalniz$ 'alone' $yalniz$-@ 'only'
   $sahin$ 'correct' $sahin$-@ 'really'
   $gercekten$ 'truth' $gercektent$-@ 'truthfully'

   (ablative)
The forms in column (i) are obtained at level 2 by the application of FSA whereas those in (ii) are derived at level 1 by attaching an abstract zero suffix and applying ISA. Similarly, \([\text{yaz}\text{en}]\) yazin 'summer' (genitive) is derived at level 2, but \([\text{yaz}\text{in}]\) yazin 'summer' (instrumental case) which means 'in the summer time' is derived at level 1 as shown in (b) above.

d) Reduplication. The reduplicated syllable is the only prefix in MST. The forms in (i) are obtained at level 1; the forms in (ii) are obtained at level 2. The examples are taken from Dobrovolsky (1976:89)

\[(1)\]
- temiz 'clean'
- tamam 'complete'
- kirmizi 'red'
- kuru 'dry'
- kati 'hard'
- eski 'old'
- carbuk 'fast'
- kar 'black'
- bog 'empty'
- mav 'blue'

These are more examples of partially reduplicated forms. The examples are taken from Zimmer (1970:160).

\[(1)\]
- tersteniz 'spotless'
- tastamam 'totally complete'
- kipkirmizi 'bright red'
- kupkuru 'bone dry'
- kastat 'stone hard'
- epeski 'very old'
- carcabuk 'very fast'
- kapkara 'coal black'
- bobog 'utterly empty'
- masmavi 'very blue'

2. Level 2 Derivations.

(a) \([\text{oda}\text{da}\text{ktler}den]\)

The following morphological rules have applied in sequence to the underived form [oda]:
- locative suffixation, relative suffixation, plural suffixation, ablative suffixation. The addition of suffixes is followed by a single application of FSA and the BEC, producing \([\text{odadakilerden}]\) 'from those who are in the room'.

(b) The suffixes -CI and -IG are nominal suffixes. They may appear in the order -CI-IG:

\([\text{milliyet}ci]\text{lik}\]

The two suffixes are attached in the order indicated at level 2. FSA applies after suffixation, stressing the final syllable as shown below, and the BEC erases internal brackets:

\([\text{milliyetcilik}]\) 'nationalism'

3. Level 1 and Level 2 Derivations.

At level 1, the suffixes -E and -IG are attached, in that order, to the underived form [bu]. ISA applies followed by the BEC:

\([\text{bura}cik]\)

This form goes to level 2, where the suffix -DE attaches, followed by FSA, SRC and the BEC, yielding the following form:

\([\text{bdracikta}]\) 'right here'

4. Level 2 and 3 Derivations.

The form \([\text{aksam}]\) obtains stress at level 2 by FSA and continues to level 3 to pick up the suffix -leyin. The BEC applies at the end of the level:

\([\text{aksam}leyin}\]

\([\text{aksamleyin}]\) 'at evening'
5. Level 1 and 3 Derivations.

The underived lexical item [ne] acquires the suffix -re at level 1, and ISA applies, assigning it initial stress:

[ne\text{re}]

The BEC applies to erase internal brackets:

[ne\text{re}]

The output of level 1, which is a lexical item, is submitted as input to the level 2 rules. In this derivation, no suffix is attached at level 2 but the form undergoes FSA and SRC:

[ne\text{re}]

At level 3, the suffixes -li and -siniz are appended, in that order, and the BEC then applies:

[[ne\text{re}]\text{li}]
[[[ne\text{re}]\text{li}]\text{siniz}]

[nere\text{lisiniz}] 'where do you come from?'

NOTES

* In writing this paper, I have profited from the suggestions and criticisms of Louise Dix, Elan Dresher, John Jensen and Chris Miller.

1 The vowels and the consonants in suffixes are capitalized to indicate that they are archiphonemes. Vowels undergo the rule of vowel harmony. Consonants are subject to either the rule of Final Devoicing or Voicing Assimilation.

2 Unless otherwise indicated, the examples given in this paper are from Lewis (1967) and Underhill (1976).
Voicing Assimilation Rule:

[-voice]

[voice]

Final devoicing applies before the Voicing Assimilation rule, as indicated below:

Final Devoicing /kitab-dir/
Voicing Assimilation kitap-tir

4. The successive attachment of level 2 suffixes produces a cyclic assignment of final stress. On the first cycle, a monomorphemic item *oda* would be assigned final stress *oda*. FSA would reapply after every morphological operation (attachment of an affix). Then, at the end of level 2, all nonfinal stresses obtained at level 2 need to be reduced by the Stress Reduction Convention (7).

However, such an analysis would lead to an unnecessary multiplication of phonological rules. It cannot account for compound stress without a destressing rule specifically posited for compounds. Such a destressing rule is forced by the need to erase the stress on the second member of the compound.

The alternative analysis which I have adopted accounts elegantly for stress assigned to words obtained through affixation, at the same time predicting the correct assignment of stress in compounds. This alternative analysis eliminates the need for a stress reduction convention specifically for compounds.

REFERENCES


word formation (lexical morphology). Morphology is often referred to as grammar, the set of rules governing words in a language. Traditionally, grammars were based on the models of classical Latin and Greek, languages which contained a large number of endings. This is the system unit on the level of morphology much as the phoneme is on that of phonology. By definition a morpheme is the smallest unit which carries meaning. Even those with highly regular grammatical systems, like Finnish or Turkish, show variants of morphemes depending on the words to which they are attached. There are some elements which appear to straddle the interface of phonology and morphology in that they show properties of phonemes and of morphemes. Due to complex phonology and morphology of Turkish, parsing introduces some ambiguous parses. We developed a morphological disambiguator with accuracy of about 98% using averaged perceptron algorithm. We also present our efforts to build a Turkish web corpus of about 423 million words. These are followed by a set of lexical morphemes with the associated morphological features. The inflectional morphemes start with a + sign, and the derivational morphemes start with a - sign. Morphology of Turkish. This morphological ambiguity needs to be resolved for further language processing. Several approaches have been proposed for morphosyntactic tagging in inflective and agglutinative languages, e.g. [2,3,4,7,9,10,13].