

# Harmony and disharmony in Turkish

*Clemens Bennink*

## *0. Introduction*

Traditionally Turkish is thought of as the classical case of vowel harmony in its most transparent form.<sup>1</sup> According to this view, all vowels in Turkish words agree in their specification for backness, and all high vowels agree with the preceding vowel in their specification for roundness. However, closer examination of the Turkish lexicon reveals countless exceptions to these statements as became apparent from the well-known paper by Clements and Sezer (1982), from which this paper obtains much of its examples. More recent contributions to the issue were offered by Goldsmith (1990) and Van der Hulst and Van de Weyer (1991). The present paper discusses three aspects of VH in Turkish, viz. disharmonic roots, the distribution of the high unround back vowel, i.e. /I/, and the so-called ‘non-initial o/ö prohibition’. I will argue that these phenomena can be accounted for by principles that are necessary anyway, i.e. the Strict Cycle Condition (SCC) and the Obligatory Contour Principle (OCP).

## *1. Vowel Harmony (VH)*

The traditional approach to Turkish VH is as follows; if the first syllable of a word contains a back vowel, subsequent syllables of the word must also contain back vowels; if the first syllable of a word contains a front vowel, subsequent vowels must be front. This is illustrated by the genitive suffixes below:

(1)	nom.sg	gen.sg	nom.pl	gen.pl
‘house’	ev	ev-in	ev-ler	ev-ler-in
‘rope’	ip	ip-in	ip-ler	ip-ler-in
‘girl’	kız	kız-ın	kız-lar	kız-lar-ın
‘stalk’	sap	sap-ın	sap-lar	sap-lar-ın

---

<sup>1</sup> Thanks to Ben Hermans and Marc van Oostendorp for help and discussion.

In addition, Turkish has a second harmony process, which concerns the feature *round*. This process is often referred to as *small vowel harmony*. It can informally be described as follows; if the first syllable in a word contains an unrounded vowel, then the subsequent vowels must be unrounded as well. If on the other hand the first syllable of a word contains a rounded vowel, then the subsequent vowels must either be rounded if they are high, i.e. /u, ü/ or unrounded if they are low, i.e. /a, e/.

(2)	nom.sg	gen.sg	nom.pl	gen.pl
'rose'	gül	gül-ün	gül-ler	gül-ler-in
'stamp'	pul	pul-un	pul-lar	pul-lar-ın
'village'	köy	köy-ün	köy-ler	köy-ler-in
'end'	son	son-un	son-lar	son-lar-ın

The examples in (1) and (2) represent the complete vowel inventory of Turkish. Following earlier work (e.g. Goldsmith 1990, Van der Hulst and Van de Weyer 1991), I assume that the vowels pattern phonologically into a set of four high and four low vowels. We thus obtain the following scheme underlyingly:

(3)	Turkish Vowel System (underlying)			
	front		back	
	non-round	round	non-round	round
high	i	ü	I	u
low	e	ö	a	o

With respect to backness we find the following opposition: the front vowels /i, ü, e, ö/ versus the back vowels /I, u, a, o/. Roundness harmony opposes the nonround vowels /i, I, e, a/ to the round vowels /ü, ö, u, o/.

2. Disharmonic roots

However, while these generalizations hold for the suffixed forms such as the genitives in (1) and (2), Clements and Sezer argue that in lexical stems these generalizations no longer hold true for modern Turkish. They claim that the harmony principles are no longer applicable within the domain of the stem itself, though they are in the affixal material. This means that back and front vowels may cooccur in lexical stems and that high vowels may disagree in

roundness with the preceding vowel. Representative examples of disharmonic stems are:

(4) Disharmonic roots

a/i	hangi	'which'	hani	'where'
a/e	kalem	'pen'	hesap	'bank account'
o/i	kalori	'calory'	fiskos	'gossip'
o/e	konser	'concert'	petrol	'petrol'
u/i	billur	'crystal'	muzip	'mischievous'
u/e	suret	'copy'	memur	'official'
a/u	marul	'lettuce'	kucak	'lap'

Notice that the /o, i/ and /u, i/ roots are disharmonic, both with respect to roundness and backness. Roots with /a, u/ and /u, i/ only conflict with respect to roundness harmony. It appears that disharmonic roots mainly contain vowels from the set {a, e, i, o, u}, and to a lesser extent vowels from the set {ö, ü}. However, the high unround back vowel, i.e. /I/, is never attested to occur disharmonically.

Clements and Sezer (1982) argue that disharmony is due to the presence of 'opaque' vowels in underlying representations. Opaque segments can be characterized as vowels that are underlyingly associated with autosegmentally represented features. In order to block VH within roots they assume that opaque vowels are 'nonundergoers', 'blockers' and 'spreaders'. These claims are illustrated in (5) for the genitive singular of the form *orkinos*:

(5)	+R	-R	+R	(R = roundness)
	[[or	kin	nos] In]	'tunny fish'
	+B	-B	+B	(B = backness)

Example (5) shows that the underlying root is fully specified for the features [round] and [back]. However, the suffix vowel is not specified for any of these features. Therefore, since it is not opaque, it will receive the values for backness and roundness from the immediately preceding root vowel, yielding *orkinosun*. Clements and Sezer's proposal, though consistent in itself, is problematic in that it, in an completely arbitrary way, makes ternary use of an essentially binary system. One time, a certain vowel is specified for '+' or '-' [round]/[back], while another time, the same feature(s) may be specified '0'. A second, and probably more serious, problem for Clements and Sezer's theory concerns its relation to current views on underspecification. Underspecification theory, as developed in the work of Kiparsky (1982), Archangeli (1984), Pulleyblank (1986) etc., claims that only part of the phonetic perceptible

structure is allowed to be underlyingly present. All predictable structure has to be left out. Thus, assuming [-back] and [+round] underlyingly, only representations of the following kind would be allowed:

- (6)
- |             |    |
|-------------|----|
| +R          | +R |
|             |    |
| [or ki nos] |    |
|             |    |
| -B          |    |

Now the question arises, how the harmonizing elements [+round] and [-back] can be prevented from spreading to the empty root vowels. Notice that application of VH would lead to the ungrammatical form *orkünös*. The dilemma is that, on the one hand, Clements and Sezer's theory makes the correct predictions (cf. (5)), though it is not in agreement with underspecification theory. On the other hand, the notion underspecification is fundamentally incompatible with the *opacity theory*, so it is hard to see how the two theories can be adjusted in order to block VH in root domains.

Goldsmith (1990) proposes the following reanalysis of Clements and Sezer's account of Turkish. He claims that words with vowels chosen entirely from the set {a, e, i, o, u} do not contain a specification for the feature [back] underlyingly, but are rather represented as follows (Goldsmith 1990: 302):

- (7)
- |          |          |          |          |       |
|----------|----------|----------|----------|-------|
| [+round] | [-round] | [+round] | [-round] |       |
|          |          |          |          |       |
| u        | i        | o        | e        | a     |
|          |          |          |          |       |
|          |          | [low]    | [low]    | [low] |

Front/backness will be predicted by a general post-lexical rule, which assigns [αback] to [αround]. In this system, the vowel *a* is neither round nor non-round, and thus neither front nor back. Since the feature [back] is not present at the underlying level, one cannot speak of vowels violating or respecting back/front harmony, therefore all combinations of vowels of the set in (7) are permitted within a stem (cf. (4)).<sup>2</sup> A severe drawback of this analysis is, that it cannot provide a similar account for violation to rounding harmony, since the underlying representation has to make reference to the feature [round].

<sup>2</sup> The three vowels of Turkish that may appear in a stem that are not of this system are {ö, ü, I}. Goldsmith assumes that there is free choice among the vowels within the stem, except that the three vowels that do not belong to the 'canonical' five vowel system may not freely appear; they appear only if they harmonize with the other vowels in the word.

Van der Hulst and Van de Weijer (1991) account for disharmony by assuming that pre-associated features are not allowed to associate to any other vowel positions within the same morpheme. This assumption yields the correct result in cases like (6), since it will block spreading.

### 3. *An alternative analysis*

Essential to my analysis is the so-called Strict Cycle Condition (SCC) (Mascaró 1976, Kiparsky 1982). This independently motivated condition claims that lexical phonological rules only apply in derived environments, or stated in the spirit of the current proposal: lexical rules have no effect within the domain of a single morpheme. Here I will employ the following very simplified version of the SCC:

#### (8) Strict Cycle Condition

Morpheme properties are invariable

In my analysis of the Turkish vowel system, vowels have the following underlying structure that has to be taken as the invariable property of the morpheme.

#### (9) Underlying inventory of Turkish vowels:<sup>3</sup>

/i/	/ü/	/I/	/u/
[+cor]	[+cor]	.	.
.	[+lab]	.	[+lab]
/e/	/ö/	/a/	/o/
[+cor]	[+cor]	.	.
.	[+lab]	.	[+lab]
[-high]	[-high]	[-high]	[-high]

From the previous discussion it became clear that VH must somehow be prevented from applying in roots. In my proposal it is not so much the case

<sup>3</sup> cf. Clements (1991):

[+labial] under C-place: defines labial consonants /p, b, m/  
under V-place: defines rounded vowels /ü, u, ö, o/  
[+cor] under C-place: defines coronal consonants /t, d, n/  
under V-place: defines front vowels /i, ü, e, ö/

that VH does not apply in roots, it is rather the effect of it, which has to be blocked. (Later on I will show that VH rules right enough apply within root domains, namely in case of the default vowel.) To see how my proposal handles the disharmonic cases, let us once more return to the form *orkinos*.

- (10) [ o r k i n o s ]  
           |               |               |  
           .               [+cor]       .  
           [+lab]       .               [+lab]  
           [-high]    ([+high])<sup>4</sup> [-high]

If no restrictions on VH apply, the feature [+lab] in the first syllable will spread to the following high front vowel. Moreover, [+cor] will spread across the word, harmonizing all vowels with respect to frontness. In short, it will be clear that such an operation would dramatically alter the entire underlying structure of the morpheme. The analysis proposed here claims that the SCC, as stated in (8), will not allow for these alternations, since the underlying structure/properties of the morpheme may not be changed. Thus, the SCC effects blocking of VH in non-derived words.

#### 4. *The distribution of the default vowel*

The proposal in the previous section suggest that in Turkish roots, all vowels may freely combine, one time creating harmonic and another time creating disharmonic roots. Yet, examination of the Turkish lexicon reveals that this conclusion is not entirely true. Disharmonic roots mainly contain vowels from the set {a, e, i, o, u}, cf. (4), and to a lesser extent vowels from the set {ö, ü}. However, the high unround back vowel, i.e. /I/, never occurs disharmonically.

Clements and Sezer even claim, that disharmonic patterns including the vowels /ö, ü, I/ are totally absent, except for the occurrence of a number of stems combining /i-ü/, which violate labial harmony. Unfortunately, they do not offer an explanation for the difference between the two vowel sets, other than by stipulation, (cf. Clements and Sezer 1982: 228).

Goldsmith (1990) motivates the different behavior of the two vowel sets on the basis of the feature [front]. Contrary to /a, e, i, o, u/, whenever stems contain /ö, ü, I/, specification of this feature is required (i.e. [+front] for /ö, ü/, and [-front] for /I/). If present, this feature must spread across the word. Notice that in my framework such an account is impossible, since I only deal

<sup>4</sup> ([+high]) is the default value for the feature [high], which is inserted right at the beginning of the phonological component. For the moment I assume that this is accounted for by the Redundancy Rule Ordering Constraint (RROC) proposed by Pulleyblank (1986).

with one value for frontness, viz. [+cor]. The high unrounded back vowel /I/ is unspecified for [cor], which means that there is no feature to spread.

Van der Hulst and Van de Weyer, using unitary features, cannot adopt Goldsmith's proposal either.<sup>5</sup> Their analysis explains how all kinds of vowel sequences are possible exceptions to harmony. However, the proposed analysis also predicts configurations, which have an empty V-slot (i.e. /I/) on the surface. (F = front, R = round, L = low, V = bare V-slot)

- (11)
- |       |       |       |       |
|-------|-------|-------|-------|
| F     | F     | R     | R     |
|       |       |       |       |
| L - V | V - V | V - V | L - V |
| /e I/ | /i I/ | /u I/ | /o I/ |

The point is that these disharmonic sequences are totally lacking. In order to exclude these patterns, Van der Hulst and Van de Weyer introduce a constraint which states that disharmonic roots do not permit bare V-positions. Whenever such a configuration arises, it will be ruled out.

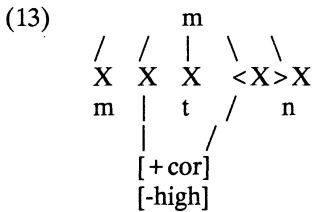
For my analysis, the patterns in (11) are problematic too, since there is no obvious reason for their absence. In order to solve this problem I will present the following analysis. In the spirit of McCarthy (1979) the status of a set of segments as a morpheme may be represented by linking them to a node called *m*. All the elements that are dominated by this node are to be considered as properties of this node. This entails that all elements on the segmental plane are accessible for the *m-node* outside that plane. In addition to this I will propose a feature that blocks interplane accessibility.<sup>6</sup> This feature is presented as <>, the so called *inaccessibility feature*. The use of this feature is limited to X-slots, which are underlyingly empty, in other words, vowels that are completely underspecified for underlying features. To understand the outlines of this idea, look at the underlying representation of *metin* 'text'.

- (12)
- |   |         |   |     |         |
|---|---------|---|-----|---------|
|   |         | m |     | /metIn/ |
| / | /       |   | \   | \       |
| X | X       | X | <X> | X       |
| m |         | t |     | n       |
|   |         |   |     |         |
|   | [+cor]  |   |     |         |
|   | [-high] |   |     |         |

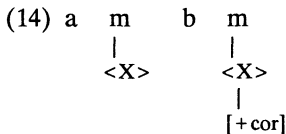
<sup>5</sup> Remember their assumption that pre-associated features ('components') cannot associate to any other vowel positions.

<sup>6</sup> cf. Hermans (forthcoming)

Notice that the segmental plane linked to  $\langle X \rangle$  is separated from the *m*-node by the feature  $\langle \rangle$ . Hence, by the very nature of the inaccessibility feature, access from the segmental plane to the dominating *m*-node is blocked. This indicates that the morpheme cannot look through the X-slot and view the segmental plane. We may say, that the plane beneath the X-slot which carries the inaccessibility feature  $\langle \rangle$  is not to be considered as a property of the morpheme. As soon as the form *metIn* enters the phonological component VH will try to take its chance by spreading the harmonizing feature  $[+cor]$  to the vowel slot at its right. We thus arrive at the following representation:



Now the question arises, how the SCC will evaluate the result of this operation. I will claim that the SCC is unable to conclude whether the properties of this morpheme have been changed or not, since the *m*-node has no access to the segmental filling of the X-slot which carries the feature  $\langle \rangle$ . In more formal terms we may say that the *m*-node cannot distinguish between the input (14a) and the output (14b).



In this fashion, the inaccessibility feature provides for a possibility to escape from the SCC, thus allowing for lexical rules to apply in underived words. I am fully aware of a severe drawback, which this analysis immediately may give rise to. Especially, when the inaccessibility theory threatens to degenerate into an *ad hoc* escape hatch, in order to avoid the strong requirements of the SCC. Therefore, I claim that the use of the inaccessibility feature must be limited to only those X-slots, which are underlyingly empty, i.e. vowels that are completely underspecified for underlying features. As a rule, these are the default vowels of a language, like the schwa in Dutch and the yer in Serbo-Croatian.



### 5. The non-initial /o, ö/ prohibition

The last topic to be discussed, concerns the wide spread claim that, the distribution of vowels {o, ö} is restricted to initial syllables of roots. Whatever its historical status may be, I will argue that this claim is only partly true for modern standard Turkish, which offers many examples of non-initial {o, ö}.

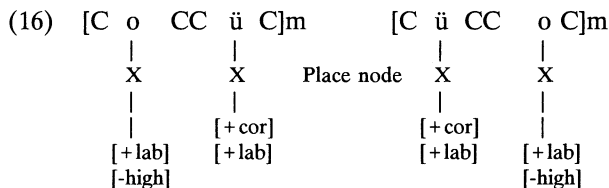
(15)	pilot	'pilot'	faktör	'factor'
	sifon	'toilet flush'	planör	'glider'
	peron	'platform'	likör	'liquor'
	metot	'method'	gresör	'lubricator'

Words of the type *pilot* also have been discussed by Clements and Sezer and Van der Hulst and Van de Weyer. However, in their frameworks, these cases are considered as normal exceptions to VH, hence, they are analogously accounted for. The examples in the second row are less frequently attested, but, as far as I can see, they are not systematically absent. Only the following sequences seem totally missing in the Turkish lexicon, namely, /o, ö/ preceded by a vowel out of the following set: {o, ö, u, ü}. The lacking of these sequences cannot be due to restrictions on harmony violations, since many of the possible orderings, like for instance /o - o/, /u - o/, /ö - ö/, /ü - ö/ etc., do not violate harmony at all. Yet, these patterns do not occur.

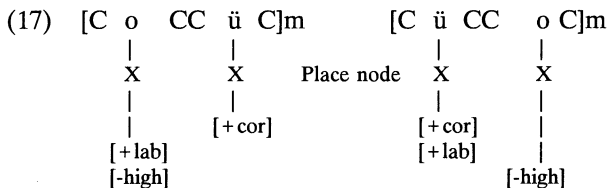
Van der Hulst and Van de Weyer suggest a general constraint, which prohibits /o, ö/ in non-initial syllables. The words in (15) are to be treated as exceptions to this constraint. I agree with Van der Hulst and Van de Weyer in that the distribution of /o, ö/ in non-initial syllables is indeed somehow restricted. However, I want to show, that this is due to independently operating principles, which govern the association of features anyway, viz. the Obligatory Contour Principle (OCP) and the SCC.

In McCarthy (1986) and Yip (1988) attention is focussed on a principle known as the OCP. It is a principle that prohibits consecutive or adjacent identical elements. Leben (1973) had observed that, in several African tonal systems, there appeared to be an effect in operation whereby, if the morphology produces a concatenation of two adjacent identical tones, the two fuse into a single tone before the tones are mapped onto their corresponding vowels. In Goldsmith (1979) it was proposed that this could be a property of

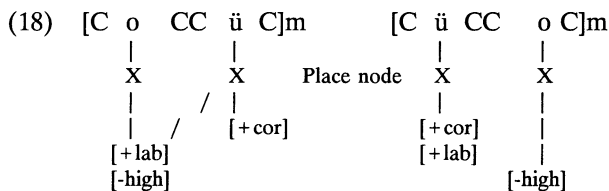
autosegmental features in general.<sup>7</sup> With this in mind we turn to the following, schematically represented, Turkish forms:



According to the OCP, these representations are ill-formed, because of the two adjacent identical elements [+lab]. OCP violations can be improved in various ways. However, with respect to Turkish, I assume that in order to correct OCP violations the second of the two identical elements has to delete. Thus we arrive at the representations below:



At this point, the derivation will be scanned for the application of VH.<sup>8</sup> Notice that after the deletion of [+lab] from the second vowel, roundness harmony gets a chance to reapply in the first form, since the initial vowel is followed by a high vowel. In the right form, the second vowel is [-high], so roundness harmony is not triggered. Therefore, the derivation terminates as below:



<sup>7</sup> With respect to *labial* I refer to OCP effects in Akkadian and Cantonese (Yip 1988). Cantonese is interesting, since it constitutes further evidence for Clements (1991). That is, in Cantonese OCP effects arise when rounded vowels clash with labial consonants. See also footnote 3.

<sup>8</sup> I assume that configurations are continually scanned for the application of rules.

At the end of the phonological component, the SCC must evaluate this representation and decide whether the underlying structure of the morpheme has been changed or not. With respect to the left form we may say that after correction of the OCP violation and the reapplication of roundness harmony, nothing has changed in the sense of the SCC. The underlying representation (16), the input, contains a low round back vowel, followed by a high round front vowel, /o/ and /ü/ resp., and so does the output in (18). However, with respect to the right hand form the derivation proceeds rather dramatically. After the OCP correction, the second vowel loses the feature [+lab], but in contrast to the first form, it gets no opportunity to receive it back, since low vowels do not trigger roundness harmony. Therefore, the form is always incorrect: underlyingly it violates OCP, while after improvement by means of [+lab] deletion it is ruled out by the SCC.

If this approach is correct, the prediction follows that low round vowels, i.e. /o, ö/, can only follow non-round vowels. Examination of the Turkish lexicon indeed indicates that this prediction is borne out:

- |      |        |           |        |           |        |
|------|--------|-----------|--------|-----------|--------|
| (19) | oküler | 'ocular'  | likör  | 'liqueur' | *ü - o |
|      | ondüle | 'curly'   | filö   | 'phil.'   | *ü - ö |
|      | formül | 'formula' | elöpen | 'lizzard' | *ö - ö |

## 6. Conclusion

I claim that there are no systematic gaps in the distribution of Turkish vowels, except for the high, unrounded, back vowel, i.e. /I/, and the /o - ö/ restriction. As for this, my analysis differs from previous analyses, in that it does not discriminate between vowels from the set {i, u, e, a, o} and those from {ö, ü, I}. Analyses which do discriminate between the two sets, will have to account for the frequently attested /i - ü/ pattern, which they predict not to occur. The same holds for the words below, in which vowels from both sets combine disharmonically.

- |      |     |         |            |        |          |
|------|-----|---------|------------|--------|----------|
| (20) | ö/a | sövalye | 'knight'   | faktör | 'factor' |
|      | ö/i | kördil  | 'folish'   | likör  | 'liquer' |
|      | a/ü | anüf    | 'nose'     | türal  | 'dust'   |
|      | u/ü | güsül   | 'ablution' | cüzur  | 'root'   |
|      | o/ü | oküler  | 'ocular'   |        |          |
|      | e/ü | erüz    | 'rice'     | güzel  | 'pretty' |

The analysis based on the SCC, straightforwardly accounts for all these cases without stipulation.

At first sight, the harmonic behavior of the default vowel seemed to undermine the SCC based analysis. However, I argued that this difficulty can be solved by the inaccessibility feature (<>). This feature can be motivated on the grounds of evarious phenomena in other languages, like for instance, the unstressability of Dutch schwa and the palatalizing behavior of some schwa's in eastern Dutch dialects (cf. Nijen Twilhaar 1990)

The last issue I discussed concerned the so called 'non-initial /ö-o/ prohibition, I argued that there is in fact no reason to assume a general constraint, which limits /o-ö/ to initial syllables. The only restriction is that these vowels cannot follow other round vowels. However, I showed that this does not need to be stipulated, but rather follows from an independent principle, viz. OCP.

### References

- Archangeli, Diana (1988) *Underspecification in Yawelmani Phonology and Morphology*, PhD dissertation, MIT 1984, Garland Press, New York.
- Clements, George N. and Sezer, Engin (1982) 'Vowel and Consonant Disharmony in Turkish', in H. van der Hulst and N. Smith, eds., *The structure of phonological representations* (Part II).
- Clements, George N. (1991) 'Place of Articulation in Consonants and Vowels', *Working papers Cornell phonetics laboratory*, 77-123.
- Goldsmith, John (1979) *Autosegmental Phonology*. PhD dissertation, MIT, 1976, IUCL, Garland Press, New York.
- Goldsmith, John (1990) *Autosegmental and Metrical Phonology*, Cambridge, Massachusetts.
- Hermans, Ben (forthcoming) *The Composite Nature of Accent*, PhD dissertation.
- Hulst, Harrie van der, and Jeroen van de Weyer (1991) 'Topics in Turkish Phonology' in R. Boeschoten and L. Verhoeven, eds., *Turkish Linguistics Today*, Leiden.
- Kiparsky, Paul (1982). From Cyclic Phonology to Lexical Phonology, in H. van der Hulst and N. Smith, eds., *The structure of phonological representations* (Part II).
- Leben, William (1973) *Suprasegmental Phonology*, PhD dissertation, MIT.
- Mascaró, Joan (1976) *Catalan Phonology and the Phonological Cycle*, PhD dissertation, MIT.
- McCarthy, John (1979) *Formal Problems in Semitic Phonology and Morphology*, PhD dissertation, MIT, Distributed by IULC.
- McCarthy, John (1986) 'OCP effects: gemination and antigemination', *Linguistic Analysis* 17, 207-63.
- Nijen Twilhaar, Jan (1990) *Generatieve fonologie en de studie van Oostnederlandse dialecten*, Amsterdam.
- Pulleyblank, Douglas (1986) 'Underspecification and low vowel harmony in Okpe', *Studies in African Linguistics* 17, 119-53.
- Yip, Moira (1988) 'The obligatory contour principle and phonological rules: a loss of identity', *Linguistic Inquiry* 19, 65-100.