Obligatory Contour Principle

Marc van Oostendorp

September 20, 2005

Background

- Last week, we introduced some of the basic principles of autosegmental phonology.
- According to this theory, features lead a phonological life which is to a large extent independent from that of other segments.
- This leads to four possible configurations of segments (x) and features (F, G):

(1)

(1a) is the 'ideal' structure, mapping one feature to one segment. If a language has feature F, it will allow (1a). But individual languages may also decide to allow one or more of the other structures in (1). If we are talking about tones — as we have been doing so far —, we have the following terminology: (1b) has a **floating tone**, (1c) has a **contour tone**, and (1d) has a **multiply linked tone**. We have seen instances of all of these in the analyses last week. The claim is that these are the *only permissible* structures.
- This week we will extend our insights into autosegmental analyses some more.

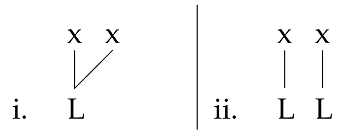
1 Multiple tones vs. multiply linked tones

Last week, we saw that Margi has three types of disyllabic words:

1. The first syllable has a low tone, the second syllable has a high tone.

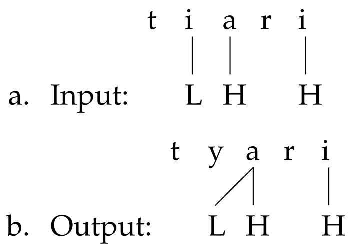
2. Both syllables have a low tone.
3. Both syllables have a high tone.

The representation of the first of these is straightforward in autosegmental terms, but for the other two, we logically speaking have two options, which I will illustrate on the low tone example:

(2)


There is one reason for assuming that the representation in (i) is the 'real' one: this allows a more uniform description of disyllabic and monosyllabic forms; recall that the latter had three tones: low, high and rising; there is no reason to assume that low toned monosyllabic stems have two (low) tones.

This reason is not very strong, but fortunately there are other arguments; and they point in the same direction. First, remember what happened to monosyllabic stems when their vowel would get lost:

(3)


And now consider the fate of bisyllabic low toned stems in the same circumstances:

(4)	làgù	làgwári	'road'
	màlà	màlári	'woman'

No rising tones are created in this case. This is much easier to understand if we assume the representation in (2a) for stems of this type than under the assumption of (2b). Under the latter representation, we would not expect any difference with monosyllabic stems: if the final vowel turns into a glide, the second low tone will go and try to find a new host on the suffix vowel, creating a contour tone in the process. But under (2a) there will be only one low tone on the stem, and this low tone does not run the risk of becoming floating, since it can still be linked to the first vowel.

We could now wonder whether there are languages which have both (2a) and (2b) in their inventory of phonological structures. It has been a claim of autosegmental phonology that this is not possible; phonological structures would be subject to the so-called *Obligatory Contour Principle* (Leben, 1973):[1]

Definition 1 (Obligatory Contour Principle (OCP)) *Adjacent identical tones are disallowed.*

The OCP allows tonal tiers like 'H', 'HL', 'L', 'LH', 'LHL', etc.; but it disallows structures like *'HH' or *'HLHLLH'. If we have two vowels in a row which are pronounced at the same pitch, there is only one option: these vowels are linked to the same tone.

2 Meeussen's Rule in Bantu

In the traditional tonology of Bantu languages, an OCP related rule is called *Meeussen's Rule* (after the Belgian Bantuist Achilles Emile Meeussen, 1912-1978). This rule can be illustrated by the following example (from KiRundi):

(5) a. nà-rá-zì-bárììrà (I-PAST-them-to sew) 'I was sewing them'
b. nà-rá-bàrììrà (I-PAST-to sew) 'I was sewing'

In (5a), the high toned tense marker *rá* and the stem *bárììrà* , which also starts with a high tone, are separated by a low tone agreement marker. Nothing happens here; we may assume that this form represents the underlying state of affairs quite faithfully. In (5b), on the other hand, the tense marker and the stem are adjacent. As a result of this, the second high tone has to go.

It is quite obviously the case that Meeussen's Rule is an OCP effect: two high tones which are adjacent are not allowed. The way to solve the OCP problem here is to turn one of the two 'bad' tones into a good tone, giving an alternation of high and low tones.

Here is another example of the same phenomenon in a different Bantu language (Shona; the data are from Odden (1980); Myers (1987); Kenstowicz (1994)):

(6)	mbwá	'dog'	né#mbwà	'with a dog'
	hóvé	'fish'	né#hòvè	'with a fish'
	mbúndúdzí	'army worms'	sé#mbùndùdzì	'like army worms'
	bàdzá	'hoe'	né#bàdzá	'with a hoe'
	bénzíbvùnzá	'inquisitive fool'	né#bènzìbvùnzá	'like an inquisitive fool'
	Fáráì	(name)	nà#Fáráì	'with Farai'

[1] Since the work of Odden (1986), it is no longer assumed that the OCP is a universal principle, but it can still be seen at work as a tendency in some languages.

What these examples show, is — among various other things — the following: (i) Meeussen's Rule applies between (certain) clitics and stems, (ii) if the clitic has a high tone and (iii) the stem starts with a high tone. Of interest are the cases in which the stem starts with more than one high-toned syllable. It turns out that in those cases, *all* of those syllables become low toned, even though for Meeussen's Rule it would be sufficient if we would only change the first one (witness what happens to forms such as *né#bàdzá*, where it also not necessary to change the second high tone of the stem.

Again, this behaviour of low toned words can be understood if we assume that again in the underlying representations two adjacent syllables pronounced at the same pitch are associated to the same tonal autosegment. If this tone has to change, all vowels attached to it will be pronounced differently:

(7) a. Input: n e h o v e — H H

b. Output: n e h o v e — H L

Interestingly, there are certain sequences of high tones which do not change together; but here, there is always an extra morpheme boundary involved. We can 'stack' clitics in Shona, leading to sequences such as:

(8) *sé#nè#hóvé* 'like with a fish'

Notice that in this case, it is only the tone of *ne* which changes. This high tone is *not* the same as the one of the stem. Therefore the latter does not automatically change with the former.

Under the assumption that Meeussen's Rule is an instance of the OCP, the latter principle actually takes two different effects in Shona:

1. It disallows sequences of the same tone in underlying forms, preferring multiply linked tones instead.
2. It disallows sequences of high tones on the surface, solving apparent problems not by spreading, but by changing one tone from high to low.

There actually is a third way in which the OCP is operative in Shona (Myers, 1987; Kenstowicz, 1994): it can also block rules from applying. This is true in particular for a rule spreading a high tone from the end of one word to the first syllable of the next word:

(9)

zvìròngó	'water pots'
zvìnà	'four'
zvìròngó zvínà	'four water pots'
Chìpó	(name)
àkàbìkà	'and then he cooked'
Chìpó ákàbìkà	'and then Chipo cooked'
ndàkáténgá	'I bought'
bàdzá	'hoe'
ndàkáténgá bàdzá	'I bought a hoe'

The last example shows that the spreading of the high tone does not occur if the second syllable of the second word already has a high tone. Spreading here would result, again, in a sequence of vowels linked to different high tones, and apparently, this is disallowed.

All in all, the OCP can thus have three effects in Shona; this is quite typical for phonological constraints (although not all of these constraints will always have all of these effects in every language):

1. It disallows certain underlying structures (by way of a **Morpheme Structure Constraint**)
2. It can **trigger** certain processes (H→L in clitic structures)
3. It can **disallow** certain processes (H spreading)

3 OCP outside of tone

The OCP gives us a good handle on extending autosegmental ideas to areas beyond tone. In many southern dialects of Dutch, the default allomorph of the diminutive suffix is *-ke* ([kə]). The following example is from Bergen Dutch (van Oostendorp, 1998):

(10) *vrouw* 'woman' - *vrouwke* 'woman-DIM' [vrɑukə]

However, if the stem ends in a velar obstruent, we find the form *-ske* ([skə]) instead (the second example also illustrates umlaut, which is irrelevant for our purposes):

(11) a. *vlieg* 'fly' - *vliegske* 'fly-DIM' [vlixskə]
b. *boek* 'book' - *buukske* 'book-DIM' [bukskə]

This can be understood as follows: bare addition of -[kə] to the stem would result in an OCP violation on the feature [velar]:

(12)
```
v l i x     k     ə
      |     |
    [vel] [vel]
```

Inserting a segment with a different place of articulation — such as coronal [s] —, solves the problem: the two segments with the 'bad place' are no longer adjacent.

A famous case of a non-tonal OCP effect is the interaction of *Lyman's Law* with *Rendaku* in (the Yamato stratum of) Japanese. The latter rule turns the second element of a compound into a voiced segment; the former expresses the condition that there is no other voiced segment elsewhere in the word (3) (Itô & Mester, 2003):

(13)	*tama*	'ball'	*teppoo*+d*ama*	'bullet'
	sono	'garden'	*hara*+z*ono*	'flower garden'
	taba	'bundle'	*satsu*+t*aba*, **satsu*-d*aba*	'wad of bills'
	sode	'sleeves'	*furi*+s*ode*, **furi*-z*ode*	'long-sleeved kimono'

Clearly, Lyman's Law — which despite its name was first discovered by Motoori Norinaga in the 18th century — could be stated as a specific instance of the OCP:

(14) *Lyman's Law* (OCP style): Avoid two voiced obstruents within the same word.

The claim is thus that Lyman's Law blocks Rendaku in Japanese in the way in which the OCP blocks high tone spreading in Shona.

There are various interesting problems connected to this. Most important among these is the issue that apparently vowels and sonorant consonants do not count for the OCP; they are, as it were, invisible as the first examples in demonstrate. The standard way of understanding this is by assuming that these segments simply do not have a link to any [±voice] feature: they are **underspecified** for that feature. Implicit in our analysis of tone above was, by the way, similarly that consonants are underspecified for tones.

4 Vowel Harmony

Another domain to which autosegmental analysis has been applied with considerable success is vowel harmony, a phenomenon that can be found in many languages of the world, albeit in different versions. In a typical vowel harmony, the set of vowels can be split up into two (or more) disjoint subsets; all the vowels within one word are taken exclusively from one subset. In Turkish, we can divide the set of vowels along the round-spread dimension as well as along the front-back dimension. The following gives a general idea of what is going on (Clements & Sezer, 1982):

(15)

	nom.sg.	*gen.sg.*	*nom.pl.*	*gen.pl.*
‘rope’	ip	ipin	ipler	iplerin
‘girl’	kız	kızın	kızlar	kızların
‘face’	yüz	yüzün	yüzler	yüzlerin
‘stamp’	pul	pulun	pullar	pulların
‘hand’	el̯	el̯in	el̯ler	el̯lerin
‘stalk’	sap	sapın	saplar	sapların

We can understand this autosegmentally by assuming that the features [±back] and [±round] can (and should) spread in Turkish:

[+back]

s a p ı n

(16) [-round]

The idea is that the phonological properties which are expressed by the harmonic features belongs to the word as a whole, and get associated to everything they can see; but they can see only those things which have a harmonic counterpart, i.e. for which the feature makes sense. Since consonants usually do not have a harmonic sister, they do not usually participate in the harmony.

Some consonants in Turkish do have a harmonic sister. They can therefore participate in the harmony as well, showing that the whole process cannot be purely harmonic (we concentrate on /k/, but similar things can be said aboud /g, l/.

(17)

-back /k̯/	+back /k/
k̯ir ‘dirt’	*kır* ‘meadows’
k̯el ‘bald’	*kul* ‘slave’
k̯ör ‘blind’	*kol* ‘arm’
dik̯ ‘upright’	*sık* ‘often’
dök̯ ‘pour’	*ok* ‘arrow’
sak̯in ‘calm’	*sıkan* ‘warning’
fak̯ir ‘poor’	*mika* ‘mica’
nek̯tar ‘nectar’	*boksit* ‘bauxite’
bol ‘abundant’	*bol̯* ‘cocktail’
kar ‘snow’	*k̯ar* ‘profit’

/k, k̯/ can also initiate harmonic behaviour themselves; to be precise on epenthetic vowels:

(18)

	careful form	*colloquial form*
'fetters'	pranga	pıranga
'prince'	prens	pirens
'test'	prova	purova
'announcer'	spiker	sipiker
'credit'	kredi	kıredi
'cruiser'	kruvazör	kuruvazör

This can also be seen in words like *kulüp* 'club', and even (given the appropriate analysis, and in certain cases) for suffixes:

(19)

	nom. sg.	*acc. sg.*
'explosion'	infil̯ak	infil̯ak̯i
'perception'	idrak	idrak̯i
'desire'	ševk	ševkı (in some dialects; older speakers)
'confirmation'	tasdik	tasdikı (in some dialects; older speakers))

Bibliography

Clements, Nick & Engin Sezer (1982). 'Vowel and Consonant Disharmony in Turkish'. In: van der Hulst, Harry & Norval Smith (eds.), *The Structure of Phonological Representations*, pp. 213–255. Dordrecht: Foris.

Itô, Junko & Armin Mester (2003). *Japanese Morphophonemics. Markedness and Word Structure*. Cambridge, Mass.: MIT Press.

Kenstowicz, Michael (1994). *Phonology in Generative Grammar*. Oxford, UK: Blackwell.

Leben, William (1973). *Suprasegmental phonology*. Ph.D. thesis, MIT, Cambridge, MA.

Myers, Scott (1987). *Tone and the structure of words in Shona*. Ph.D. thesis, University of Massachusetts at Amherst, Amherst, MA.

Odden, David (1980). 'Associative tone in Shona'. *Journal of Linguistic Research*, 1: 37–51.

——— (1986). 'On the role of the Obligatory Contour Principle in phonological theory.' *Language*, 62: 353—383.

van Oostendorp, Marc (1998). 'Bergen op Zoom Diminutives in Phonological Theory'. In: Barbiers, Sjef, Johan Rooryck, & Jeroen van de Weijer (eds.), *Small words in the big picture; Squibs for Hans Bennis*. Leiden: HIL. URL `http://www.vanoostendorp.nl/fonologie/bergen.html`.

Exercise 2

Beschouw de volgende Latijnse vormen (die bestaan uit telkens een verschillende nominale stam plus een adjectiviserend suffix), en beschrijf zo precies mogelijk wat er hier aan de hand is.

(20)
a. nav+alis 'scheeps-'
b. milit+aris 'militair'
c. mort+alis 'dodelijk'
d. sol+aris 'zonne-'
e. flor+alis 'bloem-'
f. voc+alis 'vocaal'
g. litor+alis 'kust-'