The OCP and Dissimilation in Optimality Theory*

Borim Lee
(Wonkwang University)

1. Introduction

Under the autosegmental framework, assimilation has been anayzed as spreading and dissimilation as delinking followed by (default) fill-in. It has been argued that dissimilation is triggered by the Obligatory Contour Principle, which prohibits adjacent identical elements at the melodic level (McCarthy 1986). In the traditional framework, when an OCP violation is met in the structural description of a rule, one of the identical specifications is delinked and the opposite value of the deleted feature is inserted. This process is described specifically in the rule of dissimilation whereas the OCP only functions as a background motivator. Also when dissimilation is dealt with as an autosegmental rule, there has been controversies concerning the default status of the feature fill-in rules (Odden 1987, Cohn 1992). In this paper I examine and analyze two cases of sonorant dissimilation in the newly developed Optimality theory framework and show that dissimilation phenomena are captured nicely under the constraints-based framework where OCP acts as the highest constraint. When the alternating feature values, i.e., undergoing apparent dissimilation, are properly underspecified, dissimilation is realized in the optimal output as the result of the given constraint ranking without any direct appeal to a specific rule schema. It can achieve a two-fold result: on the one hand, we can do away with a controversial dissimilation rule, and on the other hand, OCP acts as an active constraint in outputting the best

* This research has been supported by Wonkwang University Research Fund 1995.

candidate not merely existing as a backgound shadow.

The organization of this paper is as follows. In the next section the optimality theory (OT henceforth) will be briefly introduced. Under this framework, section 3 and 4 will deal with cases of sonorant dissimilation from Takelma and Sundanese respectively, which will be an Optimality-theoretic reanalysis of Lee (1994) and Cohn (1992).

2. Optimality Theory

Within the generative framework which is primarily based on the interaction of rules, contraints had received sporadic attentions as morpheme structure constraints, surface constraints or derivational constraints. They involve problems of redundancy and at best took a subsidiary role to the rule-based traditional phonology framework.

However, constraints have received a central focus in current constraint-based models (refer to Lacharite and Paradis 1993). One of the models is Optimality Theory proposed by Prince and Smolensky (1993) and developed by McCarthy and Prince (1993). In this theory there are no rules or derivations. Phonology consists of a GEN function, which produces for each input form a set of surface candidates potentially infinite, and a universal pool of constraints that are ranked from most to least important on a language-specific hierarchy. The real or optimal output is chosen by Eval(uate) function which compares candidates in terms of the constraints ranking.

(1) GEN $(input_i)$ --> $\{cand^1, cand^2, ...cand^n\}$
Eval $(\{cand^1, cand^2, ...cand^i\}) = output^{real}$

There are two important characteristics about contraints in this theory. First, contraints are in principle violable because two contraints can have incompatible requirements, and secondly, violation of constraints is minimal. Therefore, when two constraints conflict, the language fulfills the demands of the higher ranked constraint at the expense of violating the lower ranked one, hence A dominates B (A>>B).

(2)

Candidates	A	B
=> cand1		*
cand2	*	

3. Takelma sonorant dissimilation

3.1. A General Picture

Nominal stems in Takelma, an American Indian language, usually conform to CVC or CVCV templates, and as in verbal stems, the vocalic melody is restricted to just one. Some noun stems can occur as absolute nouns, but the majority of nominal stems add derivative suffixes either to stand as absolute nouns or to derive nominal phrases. Sapir (S:223) notes that many absolute nouns which end in -(a)n, -(a)m, or -(a)l are cases in which stems were combined with these derivative suffixes. On the other hand, there is a morphologically distinct set of nominal suffixes which has exactly the same form as the derivative suffixes just mentioned. Sapir (S:214) calls these noun characteristic suffixes, and they occur when the absolute forms of nouns are incremented to make a phrase. For the discussion at hand, however, the distinction does not make any differences. In the following discussion, therefore, I will not distinguish between them and refer to them simply as "nominal suffixes".

The alternation of nominal suffixes involves three sonorant consonants available in the language, and the selection of a form is usually conditioned by the last consonant in the stem. The consonant inventory of this language is as follows:

(3) consonant inventory

p	t	k	k^w		
p'	t'	k'	k'^w		
p^h	t^h	k^h	k^{wh}		
	s	x			
	ts'				
	l				
m	n				
w		y			
				h	?

Consideration of the following forms indicates that the suffix *-(a)n* occurs with a variety of stem consonants excluding stem-final *l, m* or *n,* suggesting that *-(a)n* is the basic form.

(4)	pepe-n	'rushes'	(S:47)
	kak'ɛ-n	'house ladder'	
	ts'axə-n	'blue-striped lizard'	
	yut'u-n	'white duck'	(S:223)

When the stem-final consonants are /l/ or /m/, however, the nominal suffixes are *-(a)m* or *-(a)l* respectively.

(5)	heel-am	'board'	(S:46)
	ts'el-am	'hail'	
	kel-ɛm	'river'	
	xil-am	'sick, ghost'	
	yulu-m	'eagle'	(S:47)
	kulu-m	'oak'	

(6)	soom-al-	'mountain'	(S:227)
	soom	'absolute'	
	toom-al-	'testicles'	
	toom	'absolute'	
	ts'aam-al-		
cf.	tak^h-ts'aamala?	'Klamath Indian'	
	simi-	'dew'	(S:218)

The last set of data is most puzzling in that the roots with final underlying /n/ take *-(a)m* suffix and also change the root final /n/ to [l], surfaced as *..lam.*

(7)	ha-k^waal-am	'in the road'	(S:45)
	k^waan	absolute noun for 'road'	

xaal-am-(t^hk^h)	'(my) urine'	(S:227)
xaar	absolute noun for 'urine'	
tii-tau-am	'Grant's Pass'	(S:45)
over rock	(Place name)	
tan	absolute noun for 'rock'	(S:45)

The existence of the absolute forms above makes it clear that these roots have final *n* underlyingly. Thus, two *n*'s cooccurring in this morphological environment are realized as *lVm*, not **mVl*, or any other sonorant combinations.

3.2. An optimality-theoretic analysis

Summarizing the facts given above, we noted that the nominal suffix, of which the underlying form is presumably *-(a)n*, has three variants, i.e., *-(a)n, -(a)l, and -(a)m.* Also, underlying stem *n* sometimes surfaced as *l*, but never as *m*. In short, in whatever combination the sonorants are given, they surface as a combination of *m* and *l*, where the ordering of the sequences is mostly predictable. Given a class of sonorant consonants with only three members (l, m, n), *l* and *m* are the two segments whose features are maximally different, i.e., they differ in both place feature and manner feature. The other pairs (n and l) and (n and m) differ only in one feature, the manner and place respectively.

Given that the basic idea underlying the OCP is to favor distinctness of adjacent elements, I propose that the above sonorant alternation phenomena are the result of OCP constraint working as the highest ranking constraint in this language. I divide the OCP constraint into two parts, one applying to the place tier banning *COR COR representation and the other applying to the manner tier banning *NAS NAS. At this stage, ranking between these two subparts of OCP is not critical, but the *COR COR part should make a specific reference to sonorancy, for coronal segments with differing sonorancy freely cooccur, for example, *yut'u-n* 'white duck' (S:223).

The other part of the analysis need posit different underlying specifications for the nominal suffix n and the n occurrring within the stem.

(8) Underspecified underlying representations

Stem /n/ vs. Nominal suffix /n/

```
   X              X
   |              |
 [+son]         [+son]
   |
 Place
   |
  COR
```

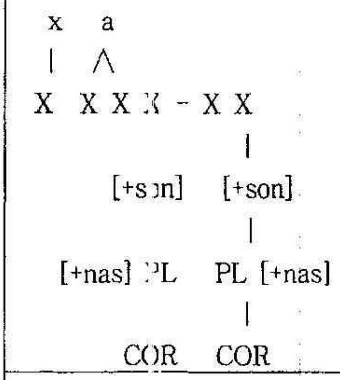
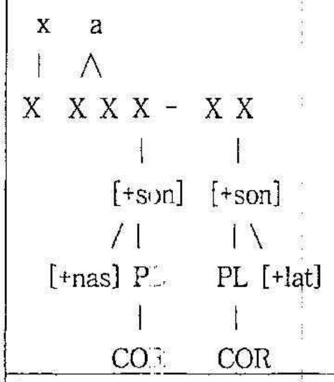
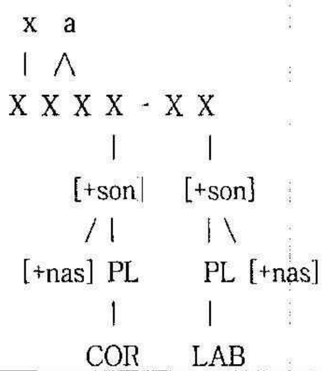
Dissimilation then is the automatic consequence in the process of obtaining the optimal output relating to the highest ranking constraints OCP(place) and OCP(manner). We do not need any specific dissimilation rules or delinking schema which merely reiterates the OCP violation and its resolution. Candidates may violate *InsertFeature constraint which belongs to Faithfulness constraint family many times, but this is to be ignored because *InsertFeature is ranked very low in the constraint hierarchy making it almost invisible.

The following tableau will make the explanation clearer.

(9) Input /xaan-an/

```
x  a
|  /\
X X X X - X  X
      |      |
   [+son]  [+son]
      |
     PL
      |
     COR
```

<table>
<tr><th>Candidates</th><th>*COR COR</th><th>*MAN MAN</th><th>*ParseFeat</th><th>*InsertFeat</th></tr>
<tr><td>=> a.
x a
| /\
X X X X - X X
| |
[+son] [+son]
/ | | \
[+lat] PL PL [+nas]
| |
COR LAB</td><td></td><td></td><td></td><td>***</td></tr>
</table>

b. 		*	*	***
c. 	*			***
d. 		*		***

<table>
<tr><td>e.
x a
| Λ
X X X X - X X
| |
[+son] [+son]
/ | | \
[+lat] PL PL [+nas]
| |
COR COR</td><td>*</td><td></td><td></td><td>***</td></tr>
<tr><td>f.
x a
| Λ
X X X X - X X
| |
[+son] [+son]
/ | | \
[+lat] PL PL [+lat]
| |
COR COR</td><td>*</td><td>*</td><td></td><td>***</td></tr>
<tr><td>g.
x a
| Λ
X X X X - X X
| |
[+son] [+son]
/ | \
[+lat] PL PL [+nas]
| |
COR COR</td><td></td><td></td><td>*</td><td>***</td></tr>
</table>

The optimal output (a) satisfies the OCP constraints. The other candidates except the last one violate the OCP once or twice, which is fatal. Even if the candidate (g) has no problem with the OCP, it violates another faithfulness contraint ParseFeat in

that it deleted - did not parse - the underlying COR place feature of the stem sonorant in the input.

These dissimilatory phenomena which the OCP contraints alone can account for can be also observed in root structures in this language.

3.3. Morpheme structure constraints relating to sonorants

Sonorant cooccurrence restrictions found in the roots are nearly identical to what we have seen in the nominal morphology. Consider the following.

(10)	a.	mal	'salmon-spear shaft'	(T:250)
		mel	'crow'	(T:250)
		lom	'cedar'	(T:249)
	b.	maan-	'count'	(T:217)
		mena	'bear'	(T:250)
	c.	*nVl, *lVn		
	d.	nanp-i-(xa)	'(his) brother's wife'	(T:251)
		luul-i-(thkh)	'(my) throat'	(T:249)

Cases of identical sonorants in (10d) suggest the possibility of double linking for identical sonorants like coocurring homorganic obstruents of this language.

(10)	a.	-t'at-(i)-	'rush'	(T:228)
	b.	ts'usu-(m)-	'make a chirping sound'	(T:232)
	c.	sas-	'stand'	(T:220)

It is not surprising to note the nonoccurrence of cases involving nonidentical coronal sonorants, i.e., l and n, as well as the abundance of forms involving maximally distinct sonorants, i.e., l and m. However, the data in (9b) present a different picture from what we have seen earlier: adjacent identical manner features, nasal, are allowed. Roots of this type are relatively rare (5 in the vocabulary of Sapir) compared to 22 roots having either mVl or lVm sequences.

In OT, the issue of morpheme structure constraints has not been dealt with in depth. The constraints in OT usually deal with morphophonemic variations. For instance, Myers (1993) discusses tonal variations in Shona under the framework of OT. His constraints, however, make a wrong prediction for underlying root level morphemes, for which Myers, in a footnote, simply notes that the given form is a root, implying that roots can be exceptions to surface constraints. This and the facts in Takelma strongly imply that some morpheme structure constraints need to be treated with a special provision. I suggest that , for Takelma, we can simply assume a doubly-linked structure of [+nas] for the roots containing two nasals, e.g., *maan* and *mena* (10c) as well as for those with totally identical nasals, e.g., *nanp-* (10d).

However, more generally, I suggest that the solution should lie in utilizing faithfulness family constraints such as PARSE respecting the underlying representations taking precedence over the contraints at hand, e.g., in the case of Takelma, OCP. Since the OCP on coronal tier, i.e., *COR COR, is consistently observed in root level as well as between morphemes, we may conclude that it was indeed a desirable move to divide the OCP into two parts where *COR COR taking precedence over *NAS NAS.

4. Sundanese dissimilation

In Sundanese, an Austronesian language, the plural marker may take the shape of *-ar- or -al-* depending on the presence of a liquid segment in the root. The plural marker usually takes the form of an infix in that it appears after the initial consonant of the roots. This phenomenon is a typical case of VC type affixation which prefixes in front of a vowel-initial root but infixes inside a consonant-initial root. Under the optimality theory this receives an explanatory analysis with two conflicting constraints, NoCoda >> Align(PrWd,L,Affix,R).

(11)

a. input /-ar-ayim/	NoCoda	Align
=> arayim		
aaryim	*	*a
ayarim		*ay
b. input /-ar-poho/		
=> paroho		*p
arpoho	*	
poharo		*poh

In Sundanese, *r* and *l* are the only [+consonantal] sonorants apart from nasals. The consonant inventory is given as follows:

(12)

p	t	c	k	
b	d	j	g	
m	n	n	N	
		s		
	l/r			
w		y		h (?)

Most roots are disyllabic, allowing maximally [CVC][CCVC]. The allomorphic variations between two plural markers are as follows:

(13)

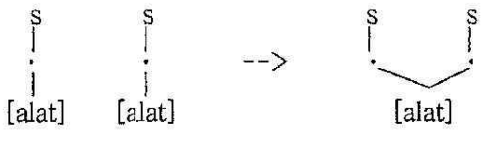
	SG	PL	
a.	rahit	r-ar-ahit	'wounded'
	riwat	r-ar-iwat	'startled'
	litik	l-al-itik	'little'
	lega	l-al-ega	'wide'
b.	curiga	c-ar-uriga	'suspicious'
	di-kirim	di-k-ar-irim	'sent' PASS

c.	gilis	g-ar-ilis	'beautiful'
	Nuliat	N-ar-uliat	'stretch'
d.	hormat	h-al-ormat	'respect'
	perceka	p-al-erceka	'handsome'
	bocor	b-al-ocor	'leaking'
	binhar	b-al-inhar	'rich'
	Numbara	N-al-umbara	'go abroad'
	siduru	s-al-iduru	'sit by a fire'
	combrek	c-al-ombrek	'cold'
	motret	m-al-otret	'take a picture'
	getol	g-ar-etol	'diligent'
	mahal	m-ar-ahal	'expensive'
	Najleh	N-ar-ajleh	'jump'
	Noplok	N-ar-oplok	'flop down'

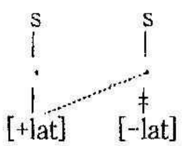
As in (13a), if a root starts with r or l, we get an assimilating infix. If the onset of the second syllable of the root is r, the result is again assimilation (13b). However, if the onset of the second syllable is l, the output is the reverse: dissimilation (13c). And the rest all involve dissmilation (13d).

Let me first present Cohn's Analysis. According to Cohn, the infix *-ar-* has an underlying r, i.e., [-lat], and she need the following rules.

(14) a. Lateral Node Merger



b. Lateral Assimilation



c. /r/-dissimilation applies to /r/ of the plural marker

```
  .         .                 .      .
  ‡         |       -->       ¦      |
[-lat]    [-lat]            [+lat] [-lat]
```

Cohn argues that OCP triggered dissimilation has its parallel forms as MSC and that dissimilation is analyzed as delinking followed by feature fill-in. She notes that in Sundanese the filled-in feature is not the default contrary to previous researches. [+lat] is filled in despite that [-lat], i.e., /r/, is less marked in this language. And she treats apparent violation of OCP as a linked structure giving evidence from Sundanese morpheme structure constraints.

Cohn's analysis has some problems. Since she follows the traditional rule schema for dissimilation, she is forced to use a more marked feature as default. Also, rules forming linked structures (rules 14a and b) are repetitive and not explanatory. She notes that in Sundanese "#/[alat]V[-alat]V.../ are rare, and when they occur they often have an alternate form of the shape /rVrV/." (1992:214) But, Cohn's rules do not directly reflect these facts.

Now, I will present an optimality-theoretic account in conjunction with underspecified structures. Since r and l are the only non-nasal sonorant consonants in Sundanese, I assume that the infix -ar-/-al- is only specified as [-nas] without any specification in laterality.

```
(15)        - a  [+son]  -
                  | \
                  |  [-nas]
                 PL
```

Since OT evaluates output candidates only, the structural condition to the dissimilation, i.e., adjacent identical features on the same tier, does not have to be present in the input. The OCP on lateral tier is the highest ranked contraint and is undominated.

To capture the initial assimilation phenomena, I propose the constraint in (16b) below which directly reflects the most noticeable MSC in Sundanese, that /#[alateral]V

[-alateral]V.../ are rare. When the onset of the second syllable is a liquid (13b and c), we get idiosyncratic results, i.e., with a root /r/ assimilation results, whereas with /l/ dissimilation. Thus, we need another contraint which resembles (16b): the constraint in (16c). There seems to be no ranking needed between (16b) and (16c).

Finally, we need a contraint against feature spreading to disallow a doubly-linked structure when it is not indispensable. Spreading, or a doubly-linked structure, is an alternative to satisfying the above mentioned specific constraints without the dominant OCP violation. Therefore, *Spreading is dominated by all the other constraints.

(16) Necessary Constraints

a. OCP on lateral tier *alat alat

b. *[s s
| |
[alat] [-alat]

c. *s s
| |
[+lat] [-lat]

d. *Spread

Although constraints 16(b) and (c) are not without redundant features, they are inevitable to cover idiosyncratic behaviors in (13b and c). In terms of contraints ranking, (16a) takes precedence over (16b and c), otherwise we get obvious OCP violations. Also, contraints (16b and c) should take precedence over *Spread; otherwise, **l-a^{r}-itik*.

In the following tableau, candidates are evaluated by the given contraints ranking and the optimal output for each candidates set is chosen.

(17)

	OCP	*[s s | | [alat] [-alat]	*s s | | [+lat] [-lat]	*Spread
(a) +l / \ => l-al-itik				*

<table>
<tr><td>+l +l
| |
l-al-itik</td><td>*</td><td></td><td></td><td></td></tr>
<tr><td>+l -l
| |
l-ar-itik</td><td></td><td>*</td><td>*</td><td></td></tr>
<tr><td>(b) -l
/ \
=> c-ar-uriga</td><td></td><td></td><td></td><td>*</td></tr>
<tr><td>-l +l
| |
c-ar-uriga</td><td>*</td><td></td><td></td><td></td></tr>
<tr><td>+l -l
| |
c-al-uriga</td><td></td><td></td><td>*</td><td></td></tr>
<tr><td>(c) -l +l
| |
=> g-ar-ilis</td><td></td><td></td><td></td><td></td></tr>
<tr><td>+l +l
| |
g-al-ilis</td><td>*</td><td></td><td></td><td></td></tr>
<tr><td>+l
/ \
g-al-ilis</td><td></td><td></td><td></td><td>*</td></tr>
</table>

The real output in (a) violates *Spread at the expense of satisfying the more highly ranked contraints. The third candidate in (a) satisfies the OCP and *Spread but the two contraints on the arrangement of liquids. In (c) the candidate satisfying every constraint fully becomes the real output.

5. Conclusion

In this study, I have examined two cases of sonorant dissimilation from two unrelated languages under the framework of Optimality Theory. In both cases, dissimilatory phenomena are OCP-triggered processes, and this fact is directly captured by the constraint-based analysis where the relevant OCP contraint acts as the undominated most highly ranked constraint. When constraints replace rules, we can do away with redundant features of any dissimilation rule which is unavoidable under the rule-based generative phonology. On the other hand, when coupled with relevant underspecifications for the dissimilating affixes, Optimality analysis can avoid the controversy over the status of filled-in features as being default or not.

References

Cohn, Abigail C. (1992) "The consequences of dissimilation in Sundanese" *Phonology* 9, 199-220.

Lee, Borim (1994) "Sonorant Cooccurrence Restrictions and the OCP" *Linguistic Journal of Korea*, 19-2, 513-537.

McCarthy, John (1986) "OCP Effects: Gemination and Antigemination," *Linguistic Inquiry* 17, 207-263.

McCarthy, John and Alan Prince (1993) "Prosodic Morphology1: Constraint interaction and Satisfaction". Ms. University of Massachusetts, Amherst, and Rutgers University, New Brunswick, NJ.

Odden David (1987) Dissimilation as deletion in Chuckchi. *ESCOL* '87. 235-246.

Prince, Alan and Paul Smolensky (1992) "Optimality: Constraint Interaction in Generative Grammar," ms., Rutgers University, New Brunswick, and University of Colorado, Boulder.

Sapir, Edward (1909) *Takelma Texts*. University of Pennsylvania anthropological publications of the University Museum 2 (1). 1-263.

Sapir, Edward (1922) *The Takelma Language of Southwestern Oregon*. Handbook of American Indian Languages, ed. by Franz Boas. Bulletin of the Bureau of American Ethnology 40(2). 1-296.