THE GRADIENT OCP: EVIDENCE FROM KOREAN REDUPLICATION

Young-ran An

Stony Brook University

1. Introduction

Korean is said to be one of the languages that make use of reduplication in order to enrich the lexicon. Ideophones (including onomatopoeia), for example, are one of the inventories which benefit from this strategy, which can be grouped into two encompassing classes, i.e., total and partial reduplications.

In this paper, I will focus on the pattern of total reduplication, under which reduplicant and base are identical. When the base begins with a consonant, reduplication is perfect.[1,2]

(1)
- a. *p^{h}otoŋ*-p^{h}otoŋ ‘chubby’
- b. *mik’ɨl*-mik’ɨl ‘slippery’
- c. *p^{h}alɨt*-p^{h}alɨt ‘verdant’
- d. *pokɨl*-pokɨl ‘simmering’
- e. *tekul*-tekul ‘rolling’
- f. *t’ok*-t’ok ‘dripping; knocking; smart’

When the base begins with a vowel, the reduplicant has a consonant inserted.[3]

(2)
- a. als’oŋ-*tals’oŋ* ‘confusing’
- b. oson-*toson* ‘on good terms’
- c. oŋki-*__t__ʃoŋki* ‘densely’
- d. alok-*__t’__alok* ‘pied’
- e. ult^{h}uŋ-*__p__ult^{h}uŋ* ‘bumpy’
- f. ʎls’iku-*__t__ʃʎls’iku* ‘whoopee’
- g. olmaŋ-*__t__ʃolmaŋ* ‘all sorts of little things (in a cluster)’
- h. ali-*__k’__ali* ‘confused’

---

[1] From here on, I will italicize the portion of the reduplicant. In this case of perfect total reduplication, I decided on the reduplicant partly by resorting to such a device as reduction test (e.g. tekul-tekul → *tek*-tekul).

[2] I am using the phonemic transcription throughout the paper, not the phonetic one. In Korean, voiced obstruents are allophones of voiceless phonemes when they are between sonorants, but it is not relevant to the focus of this paper.

[3] Here, I determined the base and reduplicant through such facts that the first morpheme in *als’oŋ-tals’oŋ* is from an independent form, *alisoŋ*, that *ʌls’iku* can be used on its own without the other morpheme, *tʃʌls’iku*, and that *olmaŋ-olmaŋ* can be used for *olmaŋ-tʃolmaŋ*, while conveying the same meaning. In addition, I had recourse to the universal markedness constraint, i.e., Onset which requires a syllable onset in the unmarked forms like reduplicants.

Rebeka Campos-Astorkiza & Jon Franco (eds.), *Papers in linguistics by the BIDE generation, Anuario del Seminario de Filología Vasca «Julio de Urquijo»* XLVI-1 (2012), 33-53.

Looking through the data given here, we see that the reduplicated consonant can vary in both place and manner of articulation. We see a bilabial consonant in (2e), a dorsal segment in (2h), and a coronal consonant in the rest of the data, and an affricate in (2c, f, g) but stops in (2a, b, d, e, h).

Then the question is whether the choice of inserted consonant is predictable. I will make an attempt to provide an analysis in the remainder of the paper. My major argument is that the inserted segment is not identical to the neighboring consonants. First of all, the relevant data are instantiated in the next section. Section 3 provides a corpus-based evidence for my argument. I also furnish experimental evidence in section 4. This evidence supports the idea of the gradient OCP, and an attempt to show this gradiency is made in section 5. The final section will wrap up the paper with the theoretical implications and future directions of the study.

2. Data

The database established based on *Essence Korean Dictionary*[4] provides 343 entries of total reduplication with an inserted or substituting[5] consonant in the onset of reduplicant. Some of the examples for each consonant by place and manner of articulation are as follows:

(3)	palatal affricates (28%)	
	a. oŋki-***tʃ****oŋki*	‘densely’
	b. ʎls’iku-***tʃ****ʎls’iku*	‘whoopee’
	c. olmaŋ-***tʃ****olmaŋ*	‘all sorts of little things (in a cluster)’
	d. ile-***tʃ****ʎle*	‘one thing or another’
	e. umul-***tʃ’****umul*	‘hesitantly’
	f. umultʃ’ʎk-***tʃ****umul****tʃ****’ʎk*	‘hesitantly’
	g. ʎls’a-***tʃ****ʎls’a*	‘delightfully’
	h. kalisan-***tʃ****ilisan*	‘bewildered’
(4)	bilabial stops (27%)	
	a. ult^{h}uŋ-***p****ult^{h}uŋ*	‘bumpy’
	b. ʎtʃʎŋ-***p****ʎ****tʃ****ʎŋ*	‘rambling’
	c. ʎli-***p****ʎli*	‘silly’
	d. otoŋ-***p****otoŋ*	‘chubby’
	e. utʃil-***p****u****tʃ****il*	‘brusque’
	f. okɨl-***p****okɨl*	‘bubbling’
	g. tʃ^{h}ail-***p***^{h}*iil*	‘procrastinate’
(5)	alveolar stops (20%)	
	a. als’oŋ-***t****als’oŋ*	‘confusing’
	b. oson-***t****oson*	‘on good terms’
	c. alok-***t****alok*	‘pied’

[4] *Essence Korean Dictionary*. 2006. Paju, Korea: Minjungseorim Co.

[5] From here on, I will use “inserted” as an umbrella term for both “inserted” and “substituting” unless otherwise indicated.

d. ʌtʃuŋi-***t'ʌtʃuŋi*** ‘rabble’
e. aoŋ-***t****aoŋ* ‘squabbling’
f. otol-***tʰ****otol* ‘knotty’
g. kʌmpul-***t****ʌmpul* ‘pell-mell’

(6) bilabial nasals (9%)
a. oŋsoŋ-***m****aŋsoŋ* ‘hazy’
b. ʌli-***m****ali* ‘drowsily’
c. kinka-***m****inka* ‘obscure’
d. seŋke-***m****aŋke* ‘groundless’

(7) velar stops (7%)
a. upul-***k'****upul* ‘windingly’
b. ona-***k****ana* ‘all the time’
c. sinan-***k****onan* ‘gradually getting worse’

(8) alveolar fricatives (5%)
a. alt'ɨl-***s****alt'ɨl* ‘extremely frugal’
b. ʌlki-***s****ʌlki* ‘entangled’
c. tʃina-***s****ena* ‘always’

(9) alveolar nasals (2%)
a. tɨls'uk-***n****als'uk* ‘uneven’
b. tɨlak-***n****alak* ‘going in and out incessantly’

Examining the above data makes it possible to present some observations as in the following:

(10) Observations of the data:

(i) If the base contains /k/ and /ŋ/, the inserted consonant can be /tʃ/ or /m/, but not /k/ or /ŋ/.
(ii) If the base contains /k/ and /l/, the inserted consonant can be /tʃ/, /p/, /t/, /s/, or /n/ but not /k/ or /l/.
(iii) If the base contains /ŋ/ and /l/, the inserted consonant can be /tʃ/, /p/, or /t/, but not /ŋ/ or /l/.
(iv) If the base contains /l/ and /l/, the inserted consonant can be /tʃ/, /pʰ/, /p/, /m/, or /s/, but not /l/.
(v) If the base contains /l/ and /m/, the inserted consonant can be /tʃ'/ or /t/, but not /l/ or /m/.
(vi) If the base contains /k/ and /m̩/, the inserted consonant can be /tʃ/, but not /k/ or /m/.
(vii) If the base contains /s'/ and /l/, the inserted consonant can be /tʃ/, but not /s'/ or /l/.
(viii) If the base contains /n/ and /l/, the inserted consonant can be /tʃ/, but not /n/ or /l/
(ix) If the base contains /ŋ/ and /tʃ/, the inserted consonant can be /p/, or /t'/, but not /ŋ/ or /tʃ/.

(x) If the If the base contains /ŋ/ and /t/, the inserted consonant can be /p/, but not /ŋ/ or /t/.
(xi) If the base contains /l/ and /tʃ/, the inserted consonant can be /p/, but not /l/ or /tʃ/.
(xii) If the base contains /n/ and /s/, the inserted consonant can be /t/, but not /n/ or /s/.
(xiii) If the base contains /ŋ/ and /ŋ/, the inserted consonant can be /t/ or /m/, but not /ŋ/.
(xiv) If the base contains /l/ and /t/, the inserted consonant can be /t^h/, but not /l/ or /t/.
(xv) If the base contains /k/ and /n/, the inserted consonant can be /m/, but not /k/ or /n/.
(xvi) If the base contains /l/ and /p/, the inserted consonant can be /k'/, but not /l/ or /p/.
(xviii) If the base contains /n/ and /n/, the inserted consonant can be /k/ or /s/, but not /n/.

On the whole, a generalization that can be made about the observations is that the inserted segment is not identical to the neighboring consonants.

3. Proposal

3.1. Place Identity Avoidance

In order to formulate the generalization given in the preceding section, I, first of all, present a hypothesis, called Hypothesis 1 that the place of the inserted consonant is not identical to those of its adjacent consonants. As it turns out, the following tables show that the inserted segments tend to be non-identical to the adjacent sounds in place.

(11) Tabulation: Place

Table 1

Place factor: The epenthetic/substituting segment is labial (2/121)

inserted C (LAB)			
preceding \ following	**LAB**	**COR**	**DORS**
LAB		1	
COR		40	34
DORS	1	39	6

Table 2

Place factor: The epenthetic/substituting segment is dorsal (5/25)

inserted C (DORS)			
following / **preceding**	**LAB**	**COR**	**DORS**
LAB			
COR	5	15	2
DORS	1	2	

Table 3

Place factor: The epenthetic/substituting segment is coronal (162/197)

inserted C (COR)			
following / **preceding**	**LAB**	**COR**	**DORS**
LAB	2		5
COR	9	50	14
DORS	6	89	22

As seen in Table 1, only 2 cases out of 121 where a labial is inserted in the reduplicant have another labial in the adjacent positions, either right or left. In Table 2 where the inserted consonant is a dorsal, 5 cases out of 25 have another dorsal as a preceding or following consonant. On the other hand, Table 3 shows 162 cases out of 197 where the inserted coronal is flanked by one coronal or two. Now a question arises here. Why can a coronal be inserted next to coronals? In the next section, we look further into the data from a different perspective.

3.2. Manner Identity Avoidance

After establishing that the place feature of the inserted consonant tends not to be identical to those of the neighboring segments, I propose another hypothesis, called Hypothesis 2 which presents that if the place of the inserted consonant is identical to those of adjacent consonants, they will be distinct in manner. The following tables display that the problematic cases in Table 3 are not problematic any more. That is to say, the “coronal-coronal-coronal” cases hardly share the manner of articulation.

(12) Tabulation: Manner for “Coronal-Coronal-Coronal” in Table 3

Table 4

Manner factor: The epenthetic/substituting segment is a fricative (0/197)

inserted C (FRICATIVE)					
following / **preceding**	**STOP**	**NASAL**	**FRICATIVE**	**AFFRICATE**	**APPROX.**
STOP					1
NASAL		1			
FRICATIVE					
AFFRICATE					
APPROX.					5

Table 5

Manner factor: The epenthetic/substituting segment is a nasal (0/197)

inserted C (NASAL)					
following / **preceding**	**STOP**	**NASAL**	**FRICATIVE**	**AFFRICATE**	**APPROX.**
STOP					
NASAL					
FRICATIVE					
AFFRICATE					
APPROX.					1

Table 6

Manner factor: The epenthetic/substituting segment is an approximant (1/197)

inserted C (APPROXIMANT)					
following / **preceding**	**STOP**	**NASAL**	**FRICATIVE**	**AFFRICATE**	**APPROX.**
STOP					1
NASAL					
FRICATIVE					
AFFRICATE					
APPROX.					

Table 7

Manner factor: The epenthetic/substituting segment is an affricate (3/197)

inserted C (AFFRICATE)					
following / preceding	**STOP**	**NASAL**	**FRICATIVE**	**AFFRICATE**	**APPROX.**
STOP	2				3
NASAL		1		1	8
FRICATIVE					2
AFFRICATE				2	
APPROX.					12

Table 8

Manner factor: The epenthetic/substituting segment is a stop (4/197)

inserted C (STOP)					
following / preceding	**STOP**	**NASAL**	**FRICATIVE**	**AFFRICATE**	**APPROX.**
STOP					
NASAL			2		3
FRICATIVE					
AFFRICATE					
APPROX.	4				1

In Table 4 and 5, neither of the inserted consonants, fricative or nasal, share the manner with the neighboring consonants. When the inserted segment is an approximant, affricate, or stop, only a few cases are attested where the inserted consonant shares the same manner with the adjacent segments, as shown in Table 6, 7, and 8, respectively. We also find sequences of “dorsal-coronal-coronal” or “coronal-coronal-dorsal” rather abundant in Table 3.

(13) Tabulation: Manner for “(Dorsal)-Coronal-Coronal” in Table 3

Table 9

(6/197)

inserted \ following	STOP	NASAL	FRICATIVE	AFFRICATE	APPROX.
STOP	1			1	39
NASAL					6
FRICATIVE					4
AFFRICATE	3	3		2	27
APPROX.					3

(14) Tabulation: Manner for "Coronal-Coronal-(Dorsal)" in Table 3

Table 10

(0/197)

preceding \ inserted	STOP	NASAL	FRICATIVE	AFFRICATE	APPROX.
STOP					
NASAL	1				
FRICATIVE					
AFFRICATE	4				2
APPROX.	7				

As seen in Table 9 and 10, few or no cases are attested where the inserted coronal share the same manner with the preceding or following consonant. Therefore, we can conclude that the inserted consonant tends to be dissimilar from the adjacent consonants in place and manner. Although no serious problem is detected with regard to the argument and its evidence, it may be intriguing to investigate the manner factor on its own.

(15) Tabulation: Manner

Table 11

Manner factor: The epenthetic/substituting segment is a fricative (0/17)

inserted C (FRICATIVE)					
following / preceding	**STOP**	**NASAL**	**FRICATIVE**	**AFFRICATE**	**APPROX.**
STOP		1			4
NASAL	1	4			1
FRICATIVE					
AFFRICATE					
APPROX.		1			5

Table 12

Manner factor: The epenthetic/substituting segment is an affricate (5/97)

inserted C (AFFRICATE)					
following / preceding	**STOP**	**NASAL**	**FRICATIVE**	**AFFRICATE**	**APPROX.**
STOP	11	8		2	12
NASAL	7	10		1	26
FRICATIVE					2
AFFRICATE				2	
APPROX.	2	2			12

Table 13

Manner factor: The epenthetic/substituting segment is an approximant (4/4)

inserted C (APPROXIMANT)					
following / preceding	**STOP**	**NASAL**	**FRICATIVE**	**AFFRICATE**	**APPROX.**
STOP					3
NASAL					1
FRICATIVE					
AFFRICATE					
APPROX.					

(16) Tabulation: Place factor for Table 14

Table 14

Manner factor: The epenthetic/substituting segment is a stop (99/187)

inserted C (STOP)					
preceding \ following	**STOP**	**NASAL**	**FRICATIVE**	**AFFRICATE**	**APPROX.**
STOP	10	19	1		35
NASAL	11	16	2	3	32
FRICATIVE			2		1
AFFRICATE				6	
APPROX.	23	13	2		11

(17) Tabulation: Place for "Nasal-Nasal-Nasal" in Table 15

Table 15

Manner factor: The epenthetic/substituting segment is a nasal (23/38)

inserted C (NASAL)					
preceding \ following	**STOP**	**NASAL**	**FRICATIVE**	**AFFRICATE**	**APPROX.**
STOP		4			6
NASAL		12			3
FRICATIVE					
AFFRICATE		1			
APPROX.		3	1	1	7

Since approximants are rarely inserted into the reduplicant, the data about the approximant insertion are not critical to my argument. However, looking through the tables above, we come to pose some questions: (i) Why can a stop be inserted in a position adjacent to stops? (ii) Why can a nasal be inserted in a position next to nasals?

The ANOVA test provided in Table 16 exhibits that the distributions of the adjacent stops as apposed to the inserted stop are significantly different ($p < .05$). On the other hand, two stops before or after an approximant seldom share the same place of articulation, according to Table 17 and 18. Thus we can observe the interaction effect of place and manner. The same argument is true with the "nasal-nasal-nasal" cases. Table 19 shows that none of the nasals are identical in the place of articulation.

Table 16

One-way ANOVA: Independent variable is the place of inserted or substituting stops, and dependent variables are the place of preceding stops (Stop A) and the place of following stops (Stop C)

	Source	Sum of Squares	df	Mean Square	F	Sig.
Stop A	Between Groups	1.35	2	0.675	6.3	0.027
	Within Groups	0.75	7	0.107		
	Total	2.10	9			
Stop C	Between Groups	1.75	2	0.875	8.167	0.015
	Within Groups	0.75	7	0.107		
	Total	2.5	9			

Table 17

Places for the first two stops in a Stop-Stop-(Approximant) combination. The shaded cells are the area where we expect no or few frequencies (1/187)

inserted \ preceding	LAB	COR	DORS
LAB			
COR	2		
DORS	13	19	1

Table 18

Places for the last two stops in an (Approximant)-Stop-Stop combination. The shaded cells are the area where we expect no or few frequencies (4/187)

following \ inserted	LAB	COR	DORS
LAB		5	2
COR		4	7
DORS	4	1	

Throughout this section, the given data confirm both of my hypotheses, namely that there is a tendency for the inserted consonant to be distinct from the adjacent consonants in place and manner of articulation. With this in mind, I proceed to look at how native speakers of Korean produce the reduplicated forms with an inserted consonant.

Table 19

Places for the Nasal-Nasal-Nasal combination.
The shaded cell is the area where we expect few or no frequencies.
All the inserted nasals happen to be labial (0/38)

inserted C (LAB)			
preceding / **following**	**LAB**	**COR**	**DORS**
LAB			
COR		5	
DORS		4	3

4. Experimental Results

For a pilot study of an online reduplication task, I hypothesize that the inserted or substituted segments will tend to be non-identical to the adjacent sounds in place and manner. The participants were 13 native speakers of Korean, who are graduate students at the State University of New York at Stony Brook. They were presented with nonsense or unheard-of morphemes. The participants were requested to write down what they regard as the most natural reduplicated forms, utilizing the given portion of the word. The test sheet is provided in the Appendix and the summary of results is shown in table 20:

(18) Table 20.[6] Consonant insertion vs. consonant substitution

item	**adjacent sounds**	**inserted consonants**																	**dissimilarity (pl. & man.)**[7]
		t	tʰ	t’	s	s’	ʧ	ʧʰ	ʧ’	p	pʰ	p’	k	kʰ	k’	m	n	h	
a	ŋ, ŋ	5			3		1		1				1			1	1		√ √
b	t, k						7						1				4		√ √
c	s, s	3			4		1						3			1			* * (sss)
d	t,t	1			4				1								1		√ √
e	k, ŋ	1			1		2						3			5	1		√ √
f	m, l				7					2			1				1		√ √
g	n, k	2			2		4						2			3			√ √
h	ŋ, l	2			4		1			1		1				2	2		√ √

[6] When a participant provided two or more possible consonant insertion cases, I considered them all in this table. However, I did not count in the cases where the participants deleted the given consonant, and where they changed the vowels instead of consonants.

[7] The check marks mean the concerned segments are distinct from the flanking consonants whereas the asterisks show the identicalness of the inserted consonant with either the preceding or the following consonant.

item	adjacent sounds	inserted consonants																	dissimilarity (pl. & man.)[7]
		t	tʰ	t’	s	s’	tʃ	tʃʰ	tʃ’	p	pʰ	p’	k	kʰ	k’	m	n	h	
i	ŋ, m	9					2	1									1		√ √
j	ŋ, ŋ	3			3		1			1						2			√ √
k	ŋ, l	4			2		2			2			1			1			√ √
l	l, k	2			4		3			1			2						√ √
m	ŋ, l	4		1	3		1						1			1		1	√ √
n	l, ŋ	1			1		8						1				1		√ √
o	k, tʃ	2					2		1	1			3		1	2	1		* √ (kk)
p	l, ŋ				7		3						1			1			√ √
q	n, k	2				1	5			1						1	1		√ √
r	n, k	2				2	5						1				2		√ √
s	l, l				2		1			1	7		1	1					√ √
t	l, ŋ	2					2									7	1		√ √
u	l, k	2					1			5			2		1				√ √
v	l, s				2		2				1								√ * (ss)
w	k, t	5			1		1			1		2	2						√ * (tt)
x	ŋ, t							1		3		1					1	2	√ √
y	k, l	4			2		2		2	2	1								√ √
z	k, t	3			2		1			2			1	1				1	√ * (tt)
aa	k, l	1		1	5		1			3			2						√ √
bb	l, s	2								5		1	3						√ √
cc	k, t				1		2					6	1						√ √
dd	k, l	1			1	1	3	1				2							√ √
ee	l, s	3			3		3			3			1						√ * (ss)
ff	k, tʃ				3							4							√ √
gg	l, ŋ	1			1				2								1		√ √
hh	l, ŋ	2	1				2												√ √
ii	k, tʃ						2		2		1								√ * (tʃtʃ)
jj	l, ŋ	1			7		1												√ √
kk	ŋ, l	1			3		3			1	1		1			4			√ √
ll	ŋ, ŋ	1					1			1			9		1				√ √
mm	l, tʃ								1	1		1	4					1	√ √
nn	l, ŋ	1			4		1			2									√ √
oo	l, k				2		2					2	1			4			√ √
pp	l, ŋ	1			2			1	3						1				√ √
qq	l, n				8					2									√ √

Even if this is a rough and not-fully-fledged test which might need a more sophisticated and systematic design, we can get a general picture of how Korean speakers choose a consonant to be inserted under total reduplication.

In the first place, we will consider the consonant insertion cases in which the test examples begin without the first onset segment, and the subjects had to insert some consonant. In Table 20, these cases are indicated with bold in the *item* column. In (b) where /t, k/ were supplied, most of the subjects opted for the segment /tʃ/, avoiding any identity to the neighboring segments in place and manner. In (c) where /s, s/ was provided, many subjects chose /s/. It seems to run counter to my prediction, but we see that /t/ and /k/ are also chosen for insertion, with similar frequency. Therefore, we cannot decide which consonant would be mostly frequently used as an inserted segment. In (e) where /k, ŋ/ were given, the majority put /m/ in the reduplicant onset. This supports the idea of identity avoidance in place and manner. This result is presumed to be due to the fact that Korean has this morpheme with an inserted /m/, viz. [**m**uŋke], already in the lexicon, and it might have influenced the subjects' decision. In (f) where /m, l/ were provided, the majority chose /s/ as an inserted consonant. In (g) where /n, k/ were provided, many subjects inserted /tʃ/. In (h) where /ŋ, l/ were presented, the majority picked out /s/. We can see a near consensus in (i) where /ŋ, m/ were provided, and /t/ was the preferred consonant in the reduplicant. In (j) where /ŋ, ŋ/ were in the base, the same number of the subjects (3 people) wrote down /t/ and /s/, respectively. In (k) where /ŋ, l/ were existing consonants, /t/ was the most favored segment. In (l) where we had [l, k] in the base, /s/ was chosen by the major number of the participants. In (m) where /ŋ, l/ were provided, /t/ was the most preferred consonant. In (n) where we had /l, ŋ/ in the base, almost all the participants chose /tʃ/. In (o) where /k, tʃ/ were provided, we obtain almost the same number of /t/, /tʃ/, /k/, and /m/. Hence, we cannot decide which one is preferred. In (p) where /l, ŋ/ were provided in the base, the majority of the subjects put /s/ in the reduplicant. What is interesting is that /l, ŋ/ in (p) and /ŋ, l/ in (h) are the same in combination but different in order. However, they show the same inserted consonant. It may indicate that Korean speakers think more of combination rather than order of the base consonants in epenthesizing a consonant in the reduplicant.

What about (u) where /l, k/ were provided? The majority chose /p/ for insertion. In (w) where /k, t/ supplied, /t/ was selected by many subjects. In (z) where /k, t/ were provided, just like in the case of (w), we see the same result as in (w). That is, we have more subjects who chose /t/ than those who chose the other sounds. However, /t/ is not an absolutely preferred inserted segment because 3 subjects chose /t/, whereas 2 subjects chose /s/ and /p/, respectively. But still, we can see some consistency among the native speakers in selecting a favored consonant when given the same set of consonants in the base. In (bb) where /l, s/ were furnished in the base, the major number of the subjects put down /p/ in the reduplicant. In (ee) where we had /l, s/, the same number of /t/, /s/, /tʃ/, and /p/ resulted. In this case, then, we cannot determine which consonant is the preferred. However, via analogy with (bb) which has an identical set of base consonants, we can assume that /p/ might be the consonant that can be used most frequently, in (ee), as well. In (ii) where we had /k, tʃ/ in the base, we got only five answers. Among them, two gave /tʃ/ and the other two gave /tʃ'/. We cannot decide on a preferred inserted consonant with these meager results. Lastly, in (ll) where there were /ŋ, ŋ/ given in the base, the vast majority epenthesized /k/. The extant /ŋ/ and the inserted /k/ are pronounced in the same place, but they are distinct from each other since the former is a sonorant (nasal) and the latter is an obstruent (stop).

Now, let us look at the substitution cases. The outcome is very similar to those given above for the insertion cases. To take an example, let us have a look at (s) where /l, l/ were given in the base. In this case, the majority chose /pʰ/ for replacing the existing onset consonant. In the instances of (jj) and (nn) where there were/l, ŋ/, we gained the same result, i.e. replacement of /m/ with /s/, by most of the subjects. The abutting /l/ and /s/ are both coronal, but they are different from each other, with the former being a sonorant (approximant) and the latter an obstruent (fricative).

In sum, the results from the online experiment with native speakers supports the corpus-based analysis, showing a tendency to choose an inserted or substituting consonant that is dissimilar from the neighboring consonants in place and manner.

5. Formal Analysis

With the hypothesis being confirmed, we explore the question whether these patterns can be analyzed within a theory of grammar that assumes ranked constraints. Now that we have established via the corpus data and the experimental results that the inserted consonant in the reduplicant is differentiated from the neighboring consonants, I argue that the patterns should be explained with stochastic constraint rankings, rather than absolute ones.

First of all, we saw from the corpus data that most of the inserted consonants are all chosen from the natural class of true obstruents. This kind of outcome can be anticipated if we think of the Sonority Principle (SP), which requires that lower sonority segments be preferred in nonmoraic (onset) positions. Due to this principle, we would prefer obstruents to sonorants, and stops and affricates to fricatives in the onset position.

(19) μ/a >> μ/e,o >> μ/i,u >> μ/r,l >> μ/m,n >> μ/v,z >> μ/f,s >> μ/b,d >> μ/p,t[8]

According to this sub-hierarchy of SP, adopted from Gnanadesikan (2004) and given in (19), μ/Y means “each Y must be parsed as a mora.” Therefore, the lowest ranked segments on the hierarchy would make the best onsets.

Next, we observed from the corpus and online experiment that the inserted consonant tends to be distinct from the adjacent segments in place and manner. Hence, we need some constraint that requires identity avoidance, such as the Obligatory Contour Principle (OCP), given in (20).

(20) Obligatory Contour Principle (OCP): No identical elements next to each other.

[8] Gnanadesikan (2004) did not put /k/ and /g/ in this hierarchy, but I assume that /k/ must be categorized with /p, t/ and /g/ with /b, d/.

We can also recall the following universal hierarchy from Alderete et al. (1999) shown in (21), which was, in turn, adopted from Prince and Smolensky (1993):

(21) Place-Markedness Hierarchy: *Pl/Lab, *Pl/Dors >> *Pl/Cor

According to what we have discussed about the occurrence of consonants in the onset position of the reduplicant, we can come up with the constraint hierarchy in (22), which would be regarded as a hierarchy for obtaining an epenthesized consonant. This result is illustrated in the following tableau where we obtain an optimal output, i.e. the actual form, [oson-*toson*], with an inserted coronal stop.

(22) SP(Onset), OCP(Place/Manner) >> Place-Markedness Hierarchy

(23) Tableau for [oson-*toson*] 'on good terms'[9]

/oson-Red/	Max-IO	SP(Onset)	OCP(Pl/Man)	*Pl/Lab, *Pl/Dors	*Pl/Cor	*Pl/Phar	Dep-BR
☞ a. oson-***t**oson*			(nt, ts)[10]		*****		*
b. oson-***p**oson*				*!	****		*
c. oson-***k**oson*				*!	****		*
d. oson-***n**oson*		n!	nn, (ns)		*****		*
e. oson-***s**oson*		s!	(ns), ss		*****		*

Now, let us consider cases where bilabial and velar stops are inserted in the reduplicant.

(24) Tableau [ultʰuŋ-***b**ultʰuŋ*] 'bumpy'

/ultʰuŋ-Red/	Max-IO	SP (Onset)	OCP (Pl/Man)	*Pl/Lab, *Pl/Dors	*Pl/Cor	*Pl/Phar	Dep-BR
☞ a. ultʰuŋ-***p**ultʰuŋ*				*	****		*
☹ b. ultʰuŋ-***t**ultʰuŋ*			(tl)		*****		*
c. ultʰuŋ-***k**ultʰuŋ*			(ŋk)	*	****		*

The candidates (24b) and (24c) can go on to the next lower constraint even with the partial violations of OCP(Pl/Man). Then we would expect that (24b) should be

[9] According to the hierarchy in (22), we predict that [oson-*dʒoson*] should be the same as the actual output, [oson-*doson*], in the status of optimality. On the one hand, we might be able to exclude [oson-*dʒoson*], saying that /tʃ/ is more sonorous than /t/. On the other hand, we might have to resort to a more finely devised OCP based on the similarity between /t/ and /s, n/, and between /tʃ/ and /s, n/. At this point, I would say that they should be both optimal according to the given hierarchy of the constraints.

[10] When neighboring consonants violate the OCP partially, i.e., identical either in place or manner, I put them in parentheses and let them go on to the next constraint.

the winning candidate, contrary to fact. This means that there must be some operating force of OCP(Pl/Man) that keeps (24b) out of the running. Here, this constraint will gradiently apply to the data. Even if we have a lexical item (24a) well established in the lexicon, we can still accept (24b) because it is not completely bad and it is better than (24c). Hence, inserting /p/ is judged to be relatively better in the given environment than the sounds in the other candidates from the viewpoint of the OCP suggested here. That is, it does not mean that (24a) is absolutely the best single output among other candidates.

When it comes to the case in which we have an inserted velar stop, we can develop a tableau like in the following:

(25) Tableau for [ali-***k'****ali*] 'confused'

/ali-Red/	Max-IO	SP(Onset)	OCP(Pl/Man)	*Pl/Lab, *Pl/Dors	*Pl/Cor	*Pl/Phar	Dep-BR
☞ a. ali-***k'****ali*				*	**		*
b. ali-***p****ali*		*!		*	**		*
c. ali-***t****ali*		*!	(lt, tl)		***		*
d. ali-***k****ali*		*!		*	**		*

In (25), due to SP(Onset), we can obtain the correct output. However, we could not exclude candidates like [ali-***p'****ali*], which would have to be chosen as another optimal form according to the above constraint hierarchy. In fact, [ali-***p'****ali*] is another actual word that has the same meaning as [ali-***k'****ali*], but it is just used less frequently. This may suggest that /k'/ is just one of the probable sounds that can be epenthesized for the given input. To put it another way, it reveals a probabilistic or stochastic hierarchy of constraints imposed upon various potential output candidates.

6. Theoretical Implications and Future Directions

The analysis proposed in this paper is not without a supporting theoretical foundation. In this section, I give a brief overview of the refined viewpoint with respect to the OCP constraint. Frisch et al. (2004) argue that "the degree of co-occurrence restriction between consonants in the Arabic verbal roots depends on place of articulation, manner, and voicing features, as well as the distance between consonant pairs." (p. 218) In line with this, they point out that the traditional OCP constraint, which they call total OCP, is defective because there are some data it cannot explain in any way. They, instead, suggest a gradient OCP.

Providing evidence from many different languages with respect to the tendency to avoid repetition —local and non-local— Frisch et al. (2004) propose the formulation of phonotactic knowledge based on the idea that phonotactic acceptability is a gradient notion which is to be reflected in the patterning of lexical items of a language. In the same vein, Albright (2006) claims that "grammar itself is probabilistic and gradient." (p. 1)

Among the cross-linguistic phenomena that utilize an OCP-like identity avoidance strategy, there are Turkish emphatic reduplication, English shm-reduplication, Javanese echo-words, Cantonese language game, and adjective reduplication in the Tengxian dialect of Chinese (Wedel 1999, Yip 1993, 1996). As Wedel (1999) points out, all of these cases, including the Korean reduplication case discussed in this paper, exhibit some level of perceptual distinctiveness between base and reduplicant.[11]

However, the "feature similarity metric" employed in Frisch et al. (2004) to explain the gradiency of OCP, given in (26), does not seem to dispense with a problem.

(26) $$\text{Similarity} = \frac{\text{Shared natural classes}}{\text{Shared natural classes} + \text{Non-shared natural classes}}$$

The formula given in (26) computes the similarity among consonants due to their natural class features. Then, examples like [omok-***tʃ****omok*] and [omok-***tʃʰ****omok*] are expected to have the same frequency of occurrence because the inserted consonants (in boldtype) are equally distinct from the adjacent consonants in place and manner, which makes them as dissimilar as possible from either of the flanking consonants. However, in fact, we encounter the former much more frequently, and we hardly find the latter. Therefore, we cannot entirely resort to this computation in accounting for the gradiency of OCP in Korean reduplication.

On the other hand, native Korean speakers were inclined to choose the same kind of consonant for insertion into reduplicant, if they were given the same set of consonants in the base. This tendency implies that there must be some relevance between the native speakers' lexical knowledge and the phonotactics of the language. This kind of correlation between the lexicon and its phonotactic constraints is persuasively stated in the quotation from Frisch et al. (2004): "…Over time, functional pressures on the language have shaped the lexicon that is to be acquired by successive generations of speakers. These functional pressures influence borrowing, the creation of nonce forms, and the loss of lexical items…" (p. 218).

Recapping the paper with a remark that Korean reduplication presents another instance displaying the effects of a gradient OCP constraint, in addition to other cases like the Semitic languages, Turkish, Javanese, etc., I would like to lay an emphasis on the argument that speakers are implicitly aware of statistical patterns in the lexicon. As seen throughout the analysis, the OCP constraint was playing a pivotal role in determining the consonants inserted under reduplication. Besides, these inserted consonants can be said to be chosen in accordance with the speakers' innate knowledge of the lexicon.

Meanwhile, the claim that the OCP constraint is not totally but gradiently respected is considered to be meaningful and useful in pursuing other kinds of phonological research. At least, it seems to fit in with our intuition and observation that

[11] Yip (1993) makes use of the constraint named *REPEAT which corresponds to the OCP. She uses that constraint because she thinks it should be more general in explaining the reduplication or echo-word data which exist at the interface of phonology and morphology. This constraint is later adopted by Wedel (1999). I employ the existing OCP constraint but our tenets are the same in that base and reduplicant avoid repetition.

all the extant data do not necessarily abide by all the constraints or rules on a completely absolute scale.

With the pilot study as a starting point, my research is heading now toward reinforcing this argument with actual data by conducting a more sophisticated and carefully designed experiment with native speakers of Korean. Furthermore, I would like to discuss some similarly behaving data from other languages, which were already mentioned in this section, in comparison with the Korean data. This discussion will establish the gradient OCP and probabilistic grammar in a more robust manner.

Appendix

Direction

Each of the following morphemes is part of some reduplicative form. You, as a native speaker of Korean, are requested to fill in each of the blanks with a copied form of the given item. When you create a reduplicant, please make sure that a segment should be different from the correspondent in the given morpheme. You can prefix or suffix the reduplicative form to the given base. Feel free to write them in Korean, should you be more comfortable with using Korean.

Instantiation

a. oŋki-***tʃoŋki***
b. als'oŋ-***tals'oŋ***
c. oson-***toson***

Questionnaires[12]

a.	-왱강-	[weŋkaŋ]
b.	-아그대-	[akte]
c.	-우세-	[use]
d.	-잣대-	[tʃatte]
e.	-웅게-	[uŋke]
f.	-얼금-	[ʌlkɨm]
g.	-오근-	[okɨn]
h.	-알강-	[alkaŋ]
i.	-엄벙-	[ʌmpʌŋ]
j.	-엉정-	[ʌŋtʃʌŋ]
k.	-얼멍-	[ʌlmʌn]
l.	-어글-	[ʌkɨl]

[12] This was given in Korean lest the words should get the subjects confused when provided in English. However, I am giving phonemic transcription here in the paper in order for them to be legible for the speakers of other languages. Some of the words are from a certain dialect of Korean, which is spoken by none of the subjects, and most of them are newly made up for the sake of this survey.

m. -얼루룽- [ʌluluŋ]
n. -옹알- [oŋal]
o. -우적- [utʃVk]
p. -옹골- [oŋkol]
q. -소곤- [sokon]
r. -새근- [sekɨn]
s. -칠레- [tʃʰile]
t. -송골- [soŋkol]
u. -우글- [ukɨl]
v. -부슬- [pusɨl]
w. -오독- [otok]
x. -포동- [potoŋ]
y. -후룩- [huluk]
z. -우둑- [utuk]
aa. -드륵- [dɨlɨk]
bb. -으슬- [ɨsɨl]
cc. -부득- [putɨk]
dd. -찌륵- [tʃ'ilɨk]
ee. -우술- [usul]
ff. -바직- [patʃik]
gg. -중얼- [tʃuŋʌl]
hh. -칭얼- [tʃʰiŋʌl]
ii. -오작- [otʃak]
jj. -몽골- [moŋkol]
kk. -대롱- [deloŋ]
ll. -웅시렁- [uŋsilʌŋ]
mm. -꾸질- [k'utʃil]
nn. -밍글- [miŋkɨl]
oo. -버글- [pʌkɨl]
pp. -종알- [tʃoŋal]
qq. -난들- [nantɨl]

References

Albright, Adam, 2006, «A grammatical account of gradient phonotactics». Presentation handout in the colloquia series of Stony Brook University, Spring 2006.

Alderete, John, Jill Beckman, Laura Benua, Amalia Gnanadesikan, John McCarthy, and Suzanne Urbanczyk, 1999, «Reduplication with Fixed Segmentism», *Linguistic Inquiry* 30-3, 327-364.

Cho, Sung-moon, 2001, «An Optimality-theoretic Analysis on the Change of the Vowel Harmony in Korean». Ms.

Coetzee, Andries and Joe Pater, 2005, «Gradient phonotactics in Muna and Optimality Theory». Ms.

Essence Korean Dictionary, 2006, Paju, Korea: Minjungseorim Co.

Frisch, Stefan A., Janet B. Pierrehumbert, and Michael B. Broe, 2004, «Similarity Avoidance and the OCP», *Natural Language & Linguistic Theory* 22, 179-228.

Gnanadesikan, Amalia, 1997, *Phonology with Ternary Scales*. Doctoral dissertation. University of Massachusetts, Amherst.

—, 2004, «Markedness and faithfulness constraints in child phonology». In Kager, Pater, and Zonneveld (eds.). *Constraints in Phonological Acquisition*. pp. 73-108.

Kager, René, 1999, *Optimality Theory*. Cambridge: Cambridge University Press.

Kang, Ongmi, 1998, «A Correspondence Theoretic Analysis on Korean Partial Reduplication», *Studies in Phonetics, Phonology, and Morphology* 4: 1-17. [In Korean]

Kim, Hyung-Soo, 2003, «A New Look at Partial Reduplication in Korean», *Studies in Phonetics, Phonology, and Morphology* 9-2, 353-380.

Kim, Sun-Hoi, 2001, «Multiple Patterns in Reduplication», *Eoneohag* 29, 109-131.

Kim, Soohee, 1996a, «Emergence of the Unmarked: Korean Partial Reduplication», *CLS* 32, Main Session: 179-189.

—, 1996b, «Emergence of the Markedness Constraint No-Lar and its split Behavior in the Reduplicative Domain», *Proceedings of WECOL*.

—, «Reduplication and the Mimetic Stratum in Korean». University of Washington. Ms.

McCarthy, John J., 1993, «A Case of Surface Constraint Violation», *Canadian Journal of Linguistics* 38-2, 169-195.

— and Alan Prince, 1994, «The Emergence of the Unmarked: Optimality in Prosodic Morphology». In M. Gonzàlez (ed.), *Proceedings of the North East Linguistic Society* 24, 333-379.

Nevins, Andrew and Bert Vaux, 2003, «Metalinguistic, Shmetalinguisitc: The Phonology of Shm-reduplication», Chicago Linguistics Society.

Park, Hansang, 2003, «Productivity of Korean Partial Reduplication», *Studies in Phonetics, Phonology, and Morphology* 9-2, 393-414.

Prince, Alan and Paul Smolensky, 1993, *Optimality Theory: Constraint Interaction in Generative Grammar*. MIT Press.

Urbanczyk, Suzanne, Forthcoming, "Reduplicative Form and the Root-Affix Asymmetry".

Wedel, Andrew, 1999, «Turkish Emphatic Reduplication». *Linguistics Research Center: Phonology at Santa Cruz*. http://repositories.cdlib.org/lrc/pasc/ 1999-8.

Yip, Moira, 1993, «Identity Avoidance in Phonology and Morphology». ROA.

—, 1996, «Repetition and its Avoidance: The Case of Javanese». ROA.