Identity Avoidance in Korean Reduplication

Young-ran An
Stony Brook University
yoan@ic.sunysb.edu

1. Introduction

Korean makes use of reduplication in order to enrich the lexicon. The language has a number of ideophones which consist of reduplicated forms and are used to express onomatopoeia. Ideophones can be grouped into two classes, total and partial reduplications. In this paper, I will focus on total reduplication. The reduplicant and base are generally identical whether the base begins with a consonant or a vowel.[1, 2] The so-called base neither stands alone in form nor has any meaning by itself.

(1) a. mik'ɨl-mik'ɨl 'slippery'
b. p^halɨt-p^halɨt 'verdant'
c. tekul-tekul 'rolling'
d. t'ok-t'ok 'dripping; knocking; smart'
e. okɨl-okɨl 'simmering'
f. allaŋ-allaŋ 'flatteringly'
g. umul-umul 'mumblingly'

[1] From here on, I will underline the portion of reduplicant. In the case of perfect total reduplication, I can determine the reduplicant partly by resorting to such a device as reduction test (e.g. tekul-tekul → tek-tekul). However, it cannot apply to all the cases of perfect total reduplication. With no sufficient evidence, the determination of the base and reduplicant for the perfect total reduplication seems to be immaterial. Albeit this, I underline the second portion of word to show that either part must be the reduplicant.

[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 occur between sonorants, but it is not relevant to the focus of this paper.

However, in some cases the initial consonant of the reduplicant is different from that of the base.

(2) a. waktal-**p**aktal ‘tumultuous’

b. saŋkɨl-**p**aŋkɨl ‘with a smiling face’

c. wʌlk^{h}ʌk-**t**ʌlk^{h}ʌk ‘chattering’

In other cases, when the base begins with a vowel, the reduplicant has an inserted consonant.

(3) a. als’oŋ-**t**als’oŋ ‘confusing’

b. ult^{h}uŋ-**p**ult^{h}uŋ ‘bumpy’

c. opul-**k**opul ‘meanderingly’

d. olmaŋ-**t**ʃolmaŋ ‘all sorts of little things (in a cluster)’

In this paper I will focus on the third category, i.e. consonant insertion cases. The reduplicative consonant can be alveolar (3a), bilabial (3b), dorsal (3c), or palatal (3d), and it can be a stop (3a, b, c) or an affricate (3d).[3]

I determined the base and reduplicant through such facts that the first morpheme in *als’oŋ-t als’oŋ* is from an independent form, *alisoŋ*, and that *olmaŋ-olmaŋ* can be used for *olmaŋ-tʃolmaŋ*, while conveying the same meaning. This pattern may not apply to all the cases of the third category above; however, there is a general tendency that the onset consonant in the base is maintained in the reduplicant. It is very unusual to skip the initial consonant of the base in the Korean reduplication process. Therefore, if the second morpheme in (3) were the base, then the

[3] We will see in the later sections that the data in (3) do not provide an exhaustive list of insertion cases. There is a wider range of choices in terms of the place and manner of articulation. In fact, all the consonants, except for /ŋ/ for an independently motivated reason, existing in the Korean phoneme inventory can appear in the reduplicant onset.

reduplicative forms would be tals'oŋ-tals'oŋ, pult^{h}uŋ-pult^{h}uŋ, kopul-kopul, tʃolmaŋ-tʃolmaŋ, rather than als'oŋ-**t**als'oŋ, ult^{h}uŋ-**p**ult^{h}uŋ, opul-**k**opul, olmaŋ-**tʃ**olmaŋ. On the other hand, I can recourse to the universal markedness constraint, i.e. ONSET which requires a syllable onset in the unmarked forms like reduplicants, and the unmarked syllable structure CV discovered in the world languages.[4] Therefore, without compelling evidence to the contrary, I assume that this reduplication is a case of epenthesis, not a case of deletion.

With regard to the set of different consonants inserted in the reduplicant, I argue that although the choice is not completely predictable, the inserted segment tends not to be identical to the neighboring consonants. I consider the relevant data and discuss the nature of inserted consonants in section 2. Section 3 discusses whether there is a default consonant for insertion. Section 4 furnishes evidence from a corpus study and a word creation experiment for an Identity Avoidance Effect, which disfavors insertion of a consonant that is identical to surrounding consonants. The experimental results address the question of whether the speakers' behavior reflects the patterns of the lexicon. The Identity Avoidance Effect can be expressed as a reflection of a gradient Obligatory Contour Principle (OCP). Sections 5 and 6 present a summary of the paper, and discussion of theoretical implications and future directions of research.

2. Data and Discussion

A search of *Essence Korean Dictionary*[5] revealed 343 entries of total reduplication with an inserted (185 entries) or replaced (158 entries) consonant in the onset of the reduplicant. The focus of this paper is the insertion cases. In order to investigate the data from the viewpoint of

[4] By the constraint ONSET, I do not refer to a mere, passive constraint which simply penalizes onsetless syllables via the constraint interaction. Rather, it is a broader and stronger concept which requires an unmarked syllable structure to contain an onset.

[5] *Essence Korean Dictionary* [eysseyns kwuke sacen]. 2006. Phacwu, Korea: Mincwungselim Co.

only phonological factors, I excluded 35 out of 185 insertion cases because there was meaning association or sound assimilation between the inserted consonant and its neighboring consonants. For instance, *ijʌl-**tʃ**^{h}ijʌl* 'Like cures like' is a set phrase originated from the Chinese characters. Thus the second portion, ***tʃ**^{h}ijʌl* 'cure fire' cannot be viewed as a pure reduplication from the first portion, *ijʌl* 'with fire.' The consonant ***tʃ**^{h}* is not inserted but the morpheme ***tʃ**^{h}i* 'cure' replaces the whole morpheme *i* 'with' in the first portion of the word. In *olɨlak-**n**elilak* 'rising and falling' *olɨ* is a stem meaning 'rise' and *neli* is another stem meaning 'fall.' They cannot be considered to constitute a genuine reduplicated form.[6] As for sound assimilation, I regarded examples like *ʌkɨt-**p**ʌkɨt* 'uneven,' *ʌsɨt-**p**isɨt* 'similar,' and *ulkɨt-**p**ulkɨt* 'colorful' as having assimilation between the last segment of the base and the inserted consonant in the reduplicant. In all the assimilation cases, the preceding consonant was /t/ and the inserted consonant was /p/ in which /t/ becomes /p/ as in /ʌkɨ**t**-**p**ʌkɨt/ → [ʌkɨ**p**-**p**ʌkɨt].

Some examples of each inserted consonant are provided in the following. The percentage given for each set of examples (4)-(10) indicates the proportion of each group of sounds out of the total of 150 items.

(4) **alveolar stops** **(29.33 %)**

a. als'oŋ-**t**als'oŋ 'confusing'

b. oson-**t**oson 'on good terms'

c. ʌlluŋ-**t**ʌlluŋ 'speckled'

d. allok-**t**allok 'pied'

e. ʌtʃuŋi-**t'**ʌtʃuŋi 'rabble'

f. otol-**t**^{h}otol 'hard and lumpy'

[6] As one of the referees pointed out, they may be compounds, rather than reduplicative forms.

(5) **bilabial stops** **(28.67 %)**

a. ult^{h}uŋ-**p**ult^{h}uŋ ‘bumpy’

b. ʌtʃʌŋ-**p**ʌtʃʌŋ ‘rambling’

c. ʌli-**p**ʌli ‘silly’

d. utʃil-**p**utʃil ‘brusque’

e. okɨl-**p**okɨl ‘bubbling’

f. otoŋ-**p^{h}**otoŋ ‘chubby’

(6) **palatal affricates** **(25.33 %)**

a. oŋki-**tʃ**oŋki ‘densely’

b. ʌls’iku-**tʃ**ʌls’iku ‘whoopee’

c. olmaŋ-**tʃ**olmaŋ ‘all sorts of little things (in a cluster)’

d. ʌls’a-**tʃ**ʌls’a ‘delightfully’

e. umul-**tʃ’**umul ‘hesitantly’

f. ollaŋ-**tʃ^{h}**ollaŋ ‘splashing gently’

(7) **velar stops** **(6 %)**

a. upul-**k**upul ‘windingly’

b. allali-**k’**allali ‘bantering sound’

(8) **alveolar fricatives** **(5.33 %)**

a. alt’ɨl-**s**alt’ɨl ‘extremely frugal’

b. ʌlki-**s**ʌlki ‘entangled’

(9) **bilabial nasals** **(2.67 %)**

a. oŋsoŋ-**m**aŋsoŋ ‘hazy’

b. ʌli-**m**ali ‘drowsily’

(10) **palatal approximants (2.67 %)**

a. illʌŋ-**j**allaŋ ‘rocking’

b. iltʃ’uk-**j**altʃ’uk ‘from side to side’

In the reduplication data of Korean, a consonant that is not present in the base appears in the reduplicant. According to Alderete *et al*. (1999), if the segments in the reduplicant are not present in the base, then they are either (i) the least marked C or V of the language; or (ii) a separate morpheme. However, the data above suggest that there is no particular favored consonant for insertion. Alternatively, we might assume that the inserted C (CI) is chosen at random from the set of all the consonants of the language. For the sake of exposition, I provide the inventory of the consonants in Korean which are potential candidates for the onset position. Unlike the other consonants, /ŋ/ cannot be placed in onset.

(11) Table 1. Consonant phoneme inventory in Korean

Place / Manner	Bilabial	Alveolar	Palatal	Velar	Glottal
Stop	p p^h p’	t t^h t’		k k^h k’	
Affricate			tʃ $tʃ^h$ tʃ’		
Nasal	m	n		ŋ	
Fricative		s s’			h
Approximant	(w)[7]	l	(j)		

[7] Some claim that Korean does not have a glide, /w/ or /j/, but a combination of /u/ and /ʌ/ and /i/ and /ʌ/, respectively. I will treat it as a glide but I do not have any intention to overrule the other claim.

Korean differentiates obstruents in terms of aspiration and tenseness. Therefore, there are three kinds of [-continuant] obstruents, i.e. lenis, aspirated, and fortis. However, for the purpose of this paper I treat them as one sound sharing the same place and manner since I only consider the two factors, place and manner in the current paper.[8] For instance, /p, p^h, p'/ will be counted as a single type of consonant. Then we have 11 consonants that can be possibly inserted into the onset of reduplicant. In the corpus study, only 7 CIs out of these 11 were attested. They were /t, p, tʃ, k, s, m, j/ in the order of highest-to-lowest frequency.

With regard to the possibility of random choice of CIs, let us suppose that all the attested consonants, /t, p, tʃ, k, s, m, j/, have an equal chance to be epenthesized. Then, for example, for any given context we expect to detect the same frequency for each possible inserted consonant. The following table shows the frequency of occurrence of each inserted consonant and its context, based on the data containing 2C-bases[9]:

[8] In my further research, I plan to sort out the data due to the place, manner, and laryngeal features, in particular with regard to the similarity calculation which I will run with my data.

[9] I chose 2C-bases to investigate the frequency of CIs because they provide the identical context for CIs in terms of the number of base consonants which are believed to have potential affect on the choice of CIs. Out of the seven CIs, /t, p, tʃ, k, s, m, j/ found in the corpus, only /t, p, tʃ/ were attested in the 2C-base data.

(12) Table 2. The frequency of each CI according to its environment

CI / Context	/t/	/p/	/tʃ/
/j/ & /ŋ/	0	1	0
/k/ & /k/	0	0	1
/k/ & /l/	6	5	0
/k/ & /m/	0	0	2
/k/ & /ŋ/	0	1	3
/k/ & /p/	0	0	1
/k/ & /t/	0	2	0
/k/ & /tʃ/	0	0	1
/l/ & /k/	0	2	2
/l/ & /m/	0	0	2
/l/ & /t/	4	0	0
/l/ & /tʃ/	0	2	0
/m/ & /k/	0	1	1
/ŋ/ & /j/	0	1	0
/n/ & /k/	0	2	0
/n/ & /l/	1	0	0
/ŋ/ & /l/	5	0	4
/n/ & /s/	2	0	0
/ŋ/ & /t/	1	5	0
/ŋ/ & /tʃ/	1	2	0
/s/ & /l/	0	0	1

What (12) demonstrates is that each candidate CI does not have an equal chance to be inserted in each context found for the 2C-bases. For instance, in the context of the preceding /k/ and the following /l/ only the CIs /t/ and /p/ show approximately the same frequency.[10]

[10] At the 8th SUNY/CUNY/NYU mini-conference (November 18, 2006), I received a comment that I could look at the frequency of all the onset consonants possible in Korean and compare the resulting statistics with the frequency pattern found in the reduplication to see whether the consonant insertion data simply repeat the overall lexical frequency. Since there has been no study about the onset frequency so far in Korean, I have yet to conduct a frequency test on the onsets in Korean. Meanwhile, Albright (2006b) provided the distribution of codas obtained from the 43,932 nouns in the Sejong project corpus (http://sejong.or.kr), which is as follows:

Concerning another possibility mentioned above, let us suppose that a certain consonant serves as a default for insertion in the reduplicant. If /t/ is a default CI, then unless there is some reason not to insert /t/, /t/ will be inserted. If /p/ is a default CI, unless there is some reason not to insert /p/, /p/ will be inserted. However, as confirmed in (12), there is no clear criterion distinguishing the cases containing a particular CI and those with other inserted consonants. For example, if we posit /t/ as a default CI, we find inserted /t/ in the environments of /k/ & /l/ and /ŋ/ & /l/, and consonants other than /t/ inserted in the environments of /l/ & /k/ and /l/ & /m/. Generalizing the environment of /t/ insertion as "between a stop and a liquid" also includes the environment of non-/t/ insertion, i.e. /l/ & /k/ and /l/ & /m/. Generalizing the environment of /t/ insertion as "between a velar stop and a liquid" still includes the environment of non-/t/ insertion, i.e. /l/ and /k/. Furthermore, even for the /t/ insertion environments, /t/ is not the only CI that is attested. For /k/ & /l/, /t/ and /p/ are approximately equally attested. For /ŋ/ & /l/, /t/ and /tʃ/ are attested in an approximate equal frequency. For example, the context of preceding /l/ and following /k/ has two cases of /p/ insertion (*okɨl-**pokɨl*** 'bubbling', *ukɨl-**pukɨl*** 'simmering') and

---

(i) Distribution of codas

a. Labials	b. Coronals	c. Velars
p 1360	t 1	k 5994
p^h 64	t^h 113	k^h 18
p' 0	t' 0	k' 6
	c 17	
	c^h 160	
	c' 0	
	s 375	
	s' 0	

The most frequent place of articulation in the final coda position of nouns is velar, followed by labial, and then by coronal. If we assume that this distributional pattern is replicated in the onset position, then we would expect velars > labials > coronals on the frequency scale. On the other hand, if we think of the OCP effect, then we may expect an inverse order of frequency pattern of the one seen in (i), i.e. coronals > labials > velars. This seems to hold roughly in the reduplicant onset frequency because we have two coronals and one labial which show the highest frequencies. However, this speculation does not help at all in investigating the relation between the coda of base and the onset of reduplicant. I will suspend a conclusion until I can gain the real facts. I plan to run the corpus analysis program that can help to see the distribution of onsets in general so that I can compare it with the distribution of onsets in reduplicant.

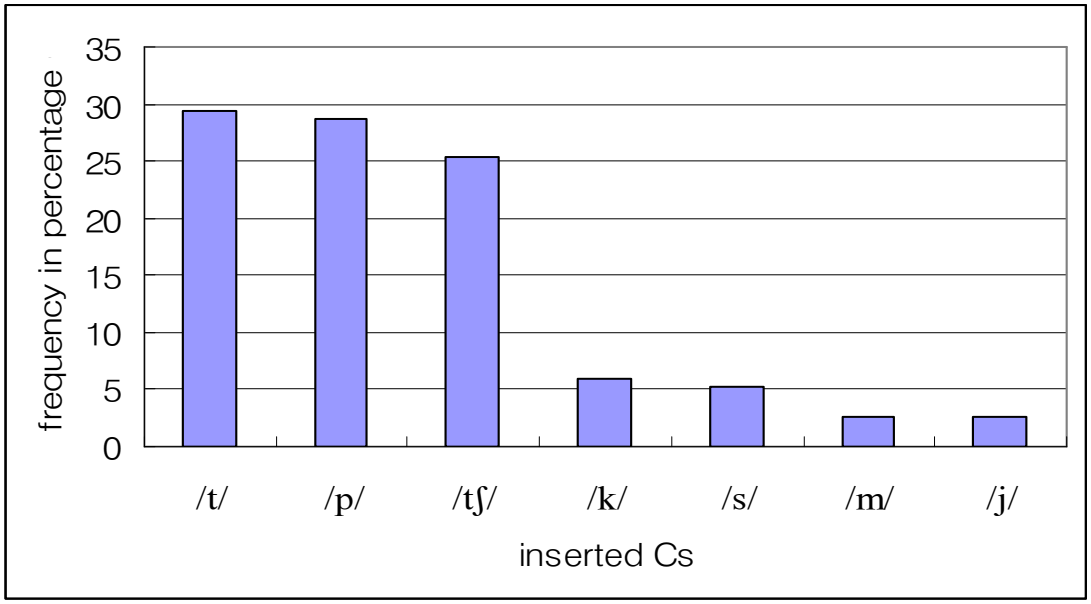
two cases of /tʃ/ insertion (*okɨl-**tʃ'**okɨl* 'wrinkled', *ukɨl-**tʃ'**ukɨl* 'crumpled'). Therefore, this renders no convincing argument for a particular default consonant as CI.[11]

Furthermore, the CI cannot be a separate morpheme in Korean. It does not mean anything on its own, having no specific grammatical function. In a nutshell, the CI is neither a default segment, nor a separate morpheme. The subsequent sections discuss the corpus and experiment results and present the hypothesis that the quality of the inserted consonant depends on the featural specifications of the neighboring base consonants.

3. Frequency of Inserted Consonants

The frequency of each CI found in the corpus study is provided in the following figure:

(13) Figure 1. The percentages of CI in the corpus data



[11] I only considered consonants here, but I do not rule out a possibility that such factors as vowels or syllables can operate in the choice of CI.

As was argued in the preceding section, there is no favored or default CI, although the three consonants /t, p, tʃ/[12] have a higher frequency than the other consonants in the corpus. As for the other classes of consonants, the sonorants /m/ and /j/ were rarely attested in the database and the sonorants /n/, /l/, and /w/ were not attested at all. This is consistent with a general preference for less sonorous consonants in the onset (cf. Selkirk 1984; Gnanadesikan 2004), according to Selkirk's universal sonority scale (p, t, k < b, d, g < f, θ < v, z, ð < s < m, n < l < r < j, w, when A is less sonorous than B in the relation of A < B). We can conclude that obstruents, which are less sonorous, are more favorable than sonorants in the onset.

For the purpose of seeing whether this frequency pattern of the corpus would be replicated in a reduplication task, an experiment was conducted to the native speakers of Korean. The participants in the experiment were 55 native speakers of Korean, ages 20s to 50s, who were recruited in Korea by the author.

The task was paper-based. The participants were presented with nonsense base morphemes, and they were asked to add a reduplicant with an inserted C to make the most natural reduplicated form with a given base. There were 3 stimuli containing one C, 19 stimuli containing two Cs, 15 stimuli containing three Cs, and 3 stimuli containing four Cs, in the given word creation task, all of which amount to 40 tokens.[13]

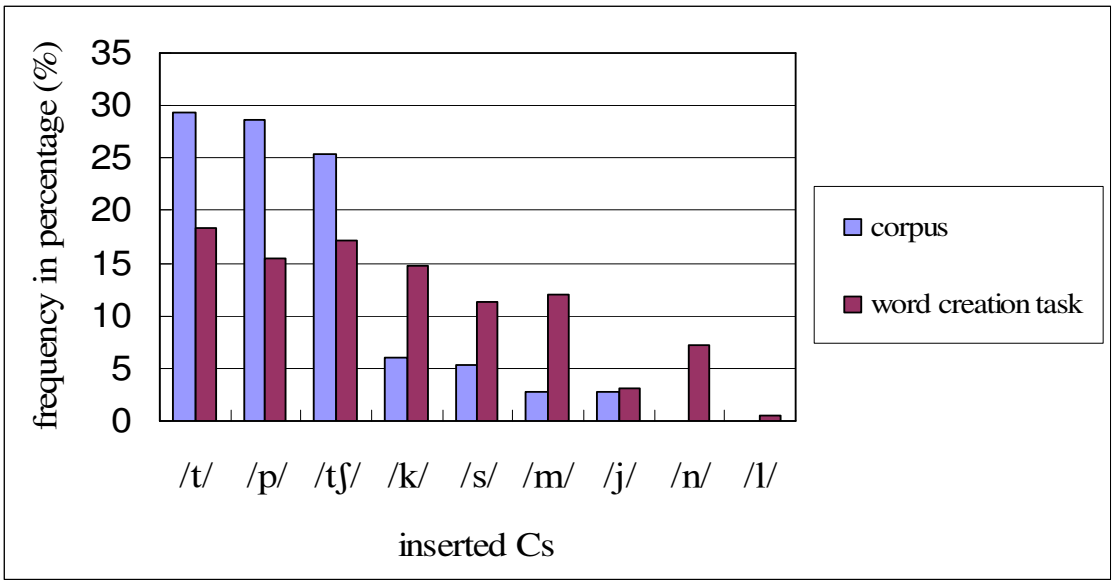
While the corpus study revealed that the three consonant classes /t, p, tʃ/ were most likely to be epenthesized in the reduplicant, the figure below shows that in the word creation task, the frequencies of CIs were more widely distributed among the various consonants, i.e. /t, p, tʃ, k, s,

[12] The consonant set of /t, p, tʃ/ includes lenis (unaspirated), fortis (tense), and aspirated variants for each of the sounds in the set.

[13] In more details, there were 2 of the base form VCV, 1 of VVC, 15 of VCVC, 3 of VCCV, 1 of CVCV, 4 of VCCVC, 9 of CVCVC, 2 of CVCCV, and 3 of CVCCVC. In my experiment, I focused on the number of consonants in the base which can affect the choice of inserted C by the participants. However, I also presume that there can be such other factors as the number of intervening Vs, the characteristics of the intervening Vs, and the kinds of Cs in the base.

m/. Therefore, we can conclude that the frequencies of CIs in the word creation task do not simply reflect the frequencies of CIs in the corpus. Furthermore, it is worth noting that these results provide even stronger evidence against the proposal that there is a default CI. Sonority is still respected; however, sonorants like /j/ and /l/ are rarely attested, with the exception of the relatively high frequency of /m/ and /n/ as CI, and other sonorants like /w/ are not attested in the experiment. Among the sonorants, nasals are less sonorous than liquids or glides. Hence the fact that nasals like /m, n/ are much more attested in the onset, abides by the tendency for a less sonorous consonant to appear in the onset. This corroborates the hypothesis that highly sonorous consonants tend not to occur as inserted consonants.

(14) Figure 2. Comparison of CI in corpus and in word creation



4. Identity Avoidance Effect

If there is no default or random choice of CIs in reduplication, then a question is raised: What

could be the criteria that guide speakers in choosing a CI in the reduplicant? One possibility is that speakers would insert a consonant that does not have the same features as other consonants of the base.

A similar restriction is found in Arabic verbal roots in which combinations of homorganic consonants in proximity are disfavored and avoided (McCarthy 1986; Frisch 2004). Arabic uses a system of 3-consonant roots, and any pair of those three consonants tends not to share the same place of articulation. The verbs like (15a, b) are not allowed since /b/ and /m/ in (15a) and /d/ and /t/ in (15b) are homorganic.[14] The actual verbs like (16a, b) do not have any pair of consonants which share the same place of articulation.

(15) a. */**b**a**m**a**h**a/

b. */**h**a**d**a**t**a/

(16) a. /**k**a**t**a**m**a/ 'to conceal; to hide'

b. /**b**a**h**a**t**a/ 'to be baffled'

Furthermore, there is evidence that Arabic speakers are implicitly aware of this avoidance of repetition or similarity. Frisch and Zawaydeh (2001) conducted an experiment using nonce verb forms that contain violations and non-violations of OCP-Place. They found that the forms with OCP-Place violations (17a) were judged by the native speakers of Arabic to be significantly less wordlike than the forms without violations (17b).

[14] In fact, the restrictions are more complicated than meets the eye. In the verbal root of $C_1VC_2VC_3$, the cases of C_1 and C_3 or C_2 and C_3 being identical in place, e.g. **l**abe**s**a 'wear,' **t**aha**n**a 'grind,' ka**s**a**d**a 'mean,' ha**s**a**l**a 'obtain,' are more tolerated than the cases in which C1 and C2 share the same place. I set aside this detailed distinction in my discussion.

(17) a. **/t**a**s**a**b**a/

b. **/t**a**h**a**f**a/

Meanwhile, the ratings varied according to the similarity of the consonant-pair stimuli as in the following:

(18) a. **/b**a**b**a**θ**a/ (identical)

b. **/θ**a**b**a**m**a/ (similar adjacent)

c. **/b**a**ʃ**a**f**a/ (similar nonadjacent)

d. **/b**a**ʔ**a**d**a/ (nonhomorganic)

The participants in the experiment rated (18a) the worst and (18d) the best on the continuum of acceptability of novel words. This shows that Arabic native speakers' judgments are affected by an OCP-Place constraint that restricts consonant co-occurrence based on similarity and proximity.

Similar patterns can be observed in diverse unrelated languages: Cantonese, English, French, Javanese, Russian, Turkish, and Yucatec Maya, to name a few.[15] The Cantonese language game "La-Mi" (Yip 1997) typically changes the first consonant of the input into /l/ and second vowel into /i/ as in (19a). However, when the input already has /l/ or /i/, a different consonant and vowel are to be used as in (19b).

(19) Cantonese La-Mi:[16] CVC > lVC CiC

a. kɔ:ŋ lɔ:ŋ kiŋ

[15] See Frisch *et al*. (2004) for more languages.
[16] See Yip (1997) for more data and discussion of Cantonese and Javanese.

	sat	lat sit		
b.	t'in	lin t'un	*lin t'in	i>u
	lat	k'at lit	*lat lit	l>k'
	lin	k'in lun	*lin lin	l>k' and i>u

In Javanese Habitual-Repetitive Reduplication (Yip 1997), the final syllable of the first half in the reduplicative form must have /a/ nucleus as in (20a), but if the input already contains /a/ as the final syllable of the base, the vowel in the final syllable of the reduplicant will have a different quality as in (20b).

(20) Javanese Habitual-Repetitive Reduplication

a.	eliŋ	elaŋ-eliŋ	'remember'	
	tuku	tuka-tuku	'buy'	
	eleʔ	elaʔ-eleʔ	'bad'	
b.	udan	udan-uden	'rain'	*udan-udan
	kumat	kumat-kumet	'have a relapse'	*kumat-kumat
	edan	edan-eden	'crazy'	*edan-edan

Turkish emphatic adjectives (Wedel 1999) are created by prefixing a CVC syllable where the initial CV is identical to the word-initial CV of the base, and the final C comes from the set /p, s, m, r/. This inserted consonant of the reduplicant tends to be non-identical to the first two consonants of the base.

(21) Turkish Emphatic Reduplication[17]

a.	kara	‘dark’	kapkara	‘pitch black’
b.	belli	‘clear’	besbelli	‘obvious’
c.	bejaz	‘white’	bembejaz	‘bright white’
d.	temiz	‘clean’	tertemiz	‘spotless’

All of these data support the idea of identity avoidance. We can find similar data from English, as well. For example, in a well-formedness judgment experiment by the native speakers of English, Hay, Pierrehumbert, and Beckman (2004) found that nonce forms like /stɹimsi/, which contains two stridents, were rated lower than similar forms like /stɹimpi/, even though words containing two strident obstruents are certainly allowed in English as in ‘space’ /speɪs/ or ‘starch’ /stɑɹtʃ/. All these data from diverse languages exhibit identity avoidance effects which in turn seem to be an emergence of the unmarked phenomenon.[18]

With this well-established foundation, I look at the Korean data in the following sections. First, I examine the effects of consonants that occur to the left of the inserted C, and then those to the right.

4.1. Left-hand effect

In this section, the distance between the CIs /t, p, tʃ/ and their context consonants will be evaluated in terms of the place and manner of articulation. I limit my focus to 2C-bases consisting of VCVC, in order to investigate the exhaustive contextual effect for the choice of CI.

[17] See Wedel (1999) for detailed discussion of the Turkish reduplication.

[18] The emergence of the unmarked follows from the crucial notions of constraint ranking and violation under domination in the Optimality Theory (OT) according to which even dominated constraints can be visibly active, becoming dominant, in such a form as reduplicant whereas they are dominated by faithfulness constraints in a base. In this vein, a constraint like Identity Avoidance must be dominated in the base but becomes dominant in the reduplicant in the presented examples.

VCVC-bases were attested most frequently in the reduplication with a CI, and amount to 40.8 % (51 items out of 125 V-initial bases with CI classes /t, p, tʃ/ in the corpus).[19]

In order to look into the effect of the local consonants on the CI, I first compare the CI with the consonant occurring to its left (CL). For example, I consider /t^{h}/ and /l/ in *otol-**t^{h}**otol* *'hard and lumpy'* and /tʃ/ and /k/ in *omok-**tʃ**omok* *'stout'* in terms of place and manner.

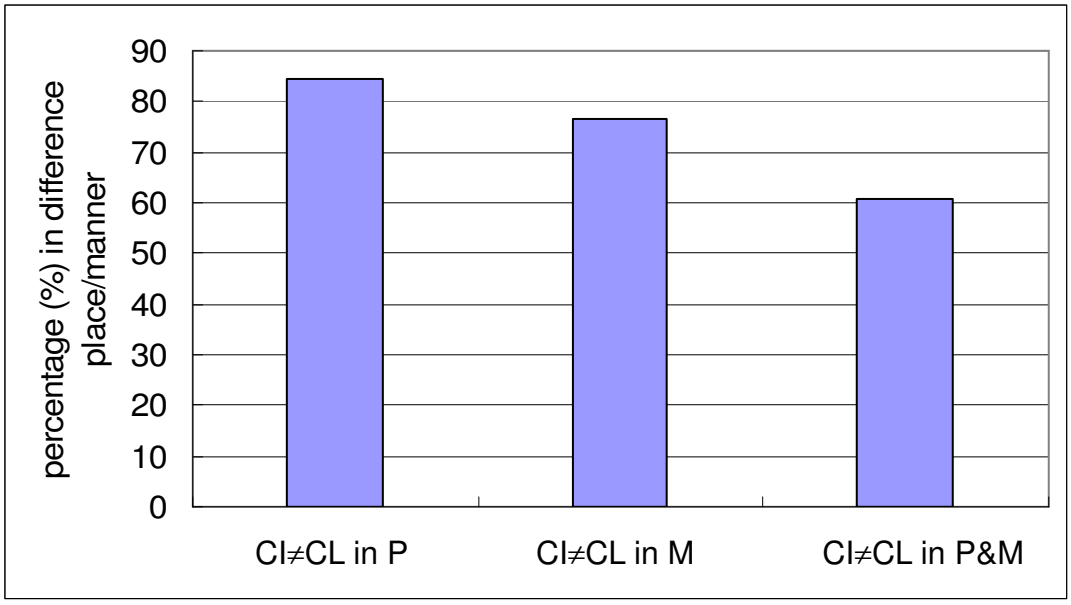
Taking account of only the VCV**C**-bases, which have a base consonant directly adjacent to the CI, we find that no CI is identical to the base-final consonant in both place and manner. The difference in place of the CI and the consonant to its left turned out to be greater than the difference in manner of those two consonants. The general point is that the CI and the consonant to its left tend to be non-identical in place and manner. In the following tables and figures, CI stands for the inserted C, CL for the consonant to the left of CI, P for place, and M for manner. When CI is distinct from CL in place or manner, I indicate it as CI≠CL in P or CI≠CL in M. When CI is different from CL in place and manner, I indicate it as CI≠CL in P&M. When CI is identical to CL in both place and manner, I indicate it as CI=CL in P&M.

(22) Table 3. VCV**C**-bases, CI=/t, p, tʃ/ from the corpus

CI vs. CL	**%**
CI≠CL in P	43/51=84.31
CI≠CL in M	39/51=76.47
CI≠CL in P&M	31/51=60.78
CI=CL in P&M	0/51=0

[19] Among the rest of the database, 3C-bases amount to 40% (50 items out of 125), 2C-bases with non-VCVC amount to 8.8% (11 items out of 125), 1C-bases amount to 8% (10 items out of 125), and 4C-bases occupy 2.4% (3 items out of 125). It may be worth looking at the data of 3C-bases in my expanded research in which I can also see if there is any distance affect between the CI and base consonants.

(23) Figure 3. VCV**C**-bases, CI=/t, p, tʃ/ from the corpus

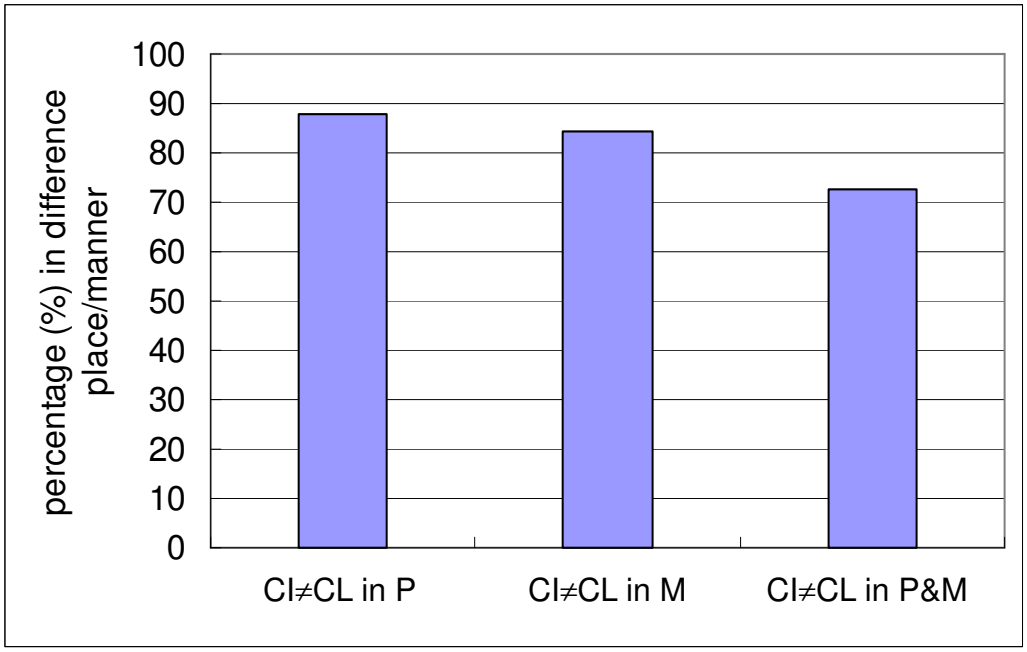


For the word creation task as well, the results are presented for only the VCV**C**-bases in the table and figure as follows:

(24) Table 4. VCV**C**-bases, CI=/t, p, tʃ/ from the word creation task

CI vs. CL	**%**
CI≠CL in P	415/472=87.92
CI≠CL in M	399/472=84.53
CI≠CL in P&M	342/472=72.46
CI=CL in P&M	0/472=0

(25) Figure 4. VCV**C**-bases, CI=/t, p, tʃ/ from the word creation task

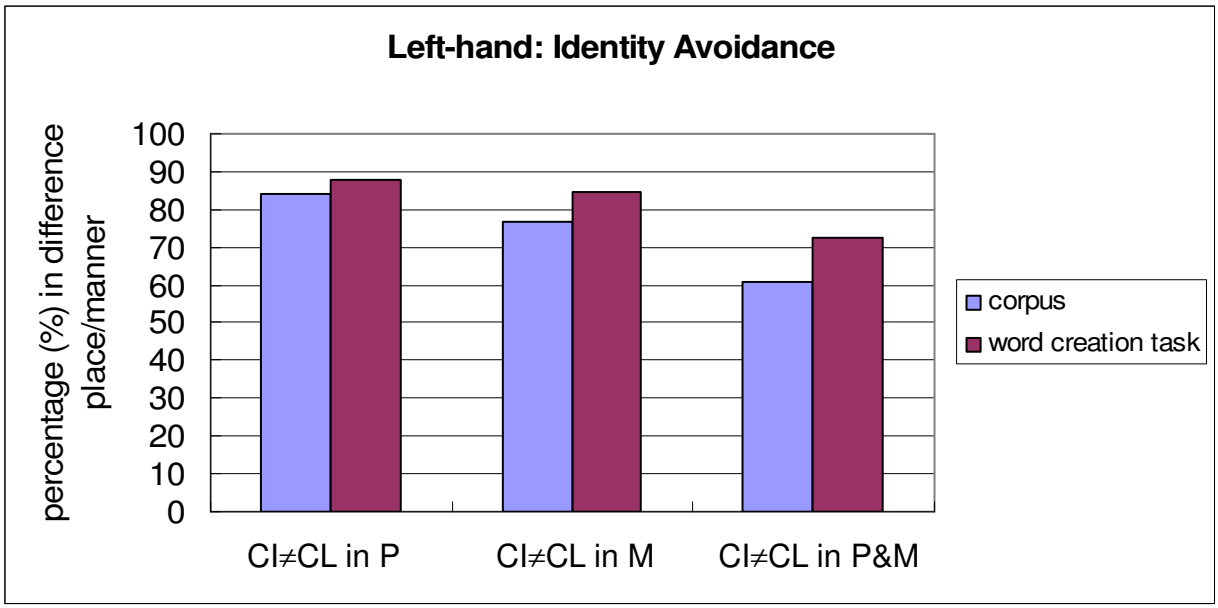


Out of the total of 40 nonsense base morphemes given to the participants, 15 items were composed of VCV**C**-bases. Among the obtained total tokens of 817,[20] 472 reduplicative forms contained the CIs /t, p, tʃ/. Since I investigate whether the results from the corpus is replicated in the experiment, I only consider the reduplicative forms which contain the CIs /t, p, tʃ/. In Table 4 we can see that there is no case in which the CI and the consonant to its left show identity in place and manner. Therefore, as in the corpus results, we can conclude that there is a tendency for the CI to be different from its left adjacent consonant in place or manner. We can also maintain the argument that the place factor plays a greater role in discriminating the CI from CL.

The following comparison figure shows that the word creation task reveals an even stronger effect in avoiding identity in place and manner between the inserted C and the consonant to its left.

[20] 15 items by the 55 participants should yield 825 reduplicated forms, but 8 items were removed since they were ill-formed or simply not answered by the participants.

(26) Figure 5. CI vs. CL



4.2. Right-hand effect

In this section, I compare the CI and the consonant to its right (CR) in terms of place and manner. For instance, I compare the place and manner of /t^{h}/ and /t/ in *otol-**t^{h}**otol* *'hard and lumpy,'* and /tʃ/ and /m/ in *omok-**tʃ**omok* *'stout.'* As in the preceding section, I look at the same V**C**VC-bases with CI = /t, p, tʃ/, for which the total number of forms is 51 items for the corpus, and 472 items for the experiment. First, the corpus outcome is given below:

(27) Table 5. V**C**VC-bases, CI=/t, p, tʃ/ from the corpus

CI vs. CR	**%**
CI≠CR in P	33/51=64.71
CI≠CR in M	34/51=66.67
CI≠CR in P&M	21/51=41.18
CI=CR in P&M	5/51=9.80

(28) Figure 6. V**C**VC-bases, CI=/t, p, tʃ/ from the corpus

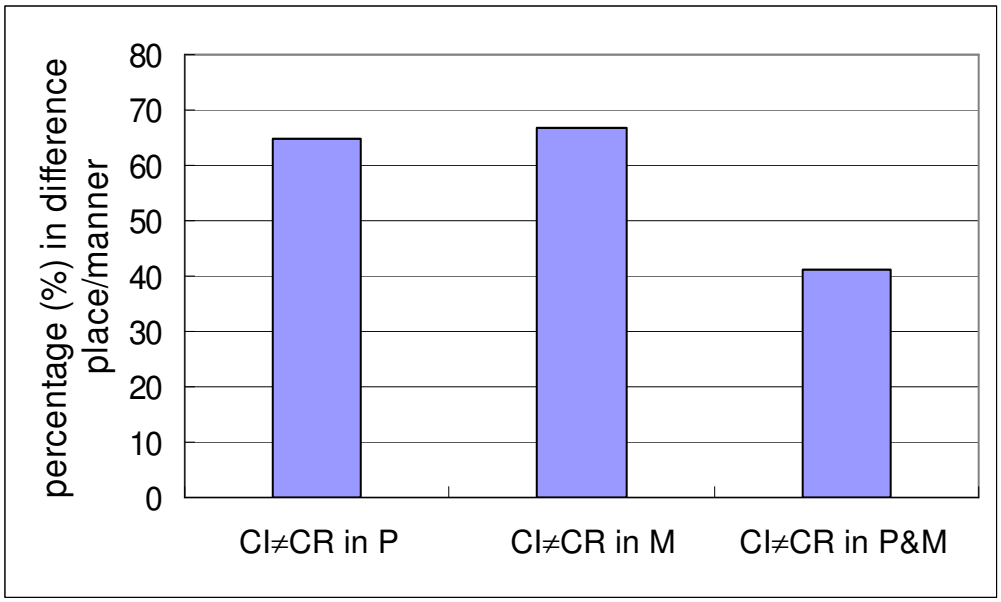


Table 5 shows that there are five pairs of the CI and the consonant to its right which share the same place and manner (9.80%). Four items have /tt^{h}/ or /t^{h}t/ (*ot^{h}ol-**t**ot^{h}ol* 'rugged,' *ut^{h}ul-**t**ut^{h}ul* 'rough,' *otol-**t^{h}**otol* 'uneven,' *utul-**t^{h}**utul* 'bumpy'). These are dissimilar in aspiration. According to Frisch *et al.*'s (2004) feature similarity metric,[21] laryngeal features should be taken into account, although I did not consider those features separately in this paper. One last sequence, /tt/ (*etoŋ-**t**etoŋ* 'childlike'), has identical sounds, and it seems to be exceptional to the suggested account. This is the only case that is not consistent with the present analysis. The overall results, however, show that there is an Identity Avoidance Effect between CI and CR. There is a greater effect of manner than of place in distinguishing the CI and CR, but the difference cannot be

[21] Frisch *et al.* (2004) calculates similarity by comparing the number of shared and unshared natural classes of two consonants, using the following equation:

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

If the identity avoidance is in operation, the greater the similarity value is, the smaller chances there are for any two segments to co-occur.

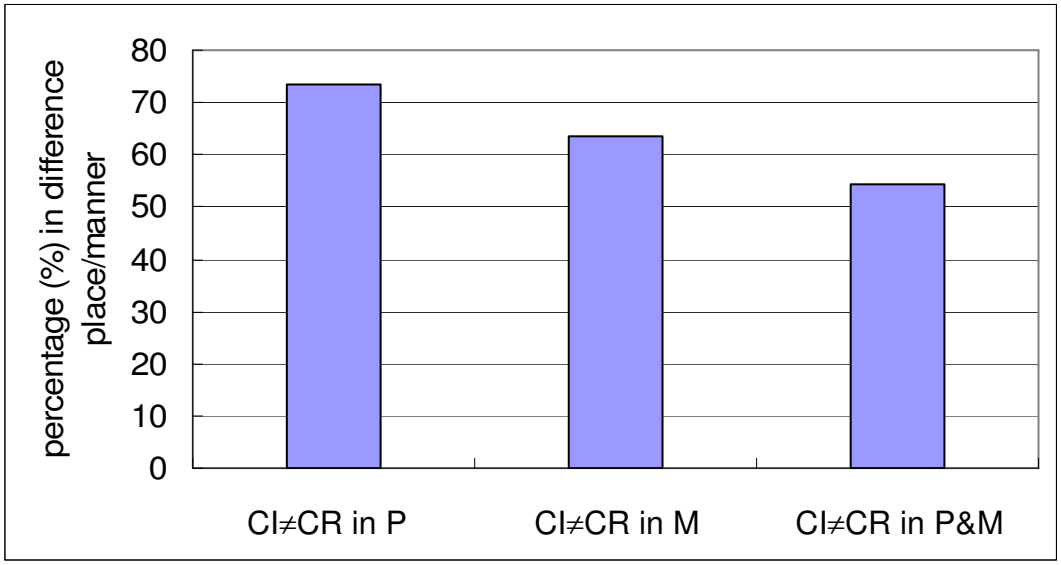
significant because there is only one case more for the manner effect than for the place effect. Therefore, we cannot claim that manner is more important in differentiating the CI and CR.

For the word creation task, I also investigated V**C**VC-bases where CI = /t, p, tʃ/:

(29) Table 6. V**C**VC-bases, CI=/t, p, tʃ/ from the word creation task

CI vs. CR	**%**
CI≠CR in P	346/472=73.31
CI≠CR in M	300/472=63.56
CI≠CR in P&M	257/472=54.45
CI=CR in P&M	83/472=17.58

(30) Figure 7. V**C**VC-bases, CI=/t, p, tʃ/ from the word creation task



As seen in the table in (29), we have 83 cases out of the total 472 items in which CI and CR are identical in place and manner. A closer examination of the given data reveals that these 83 cases consist of 9 types. Out of these 9 types, *otok-**tʰ**otok*, *atal-**tʰ**atal*, and *utʃʰil-**tʃ**utʃʰil* have /tʰt/ or /tʃtʃʰ/ which are already different in manner via the laryngeal feature. Four types out of 9 contain /tt/ or /tʃtʃ/ (*otok-**t**otok*, *atal-**t**atal*, *atʃaŋ-**tʃ**atʃaŋ*, and *otʃak-**tʃ**otʃak*). However, each of the

reduplicated forms happens to have a perfect total reduplication counterpart existing in the corpus:

(31) a. totok-totok ‘full of knobs’
b. tatal-tatal ‘stammering’
c. tʃatʃaŋ-tʃatʃaŋ ‘hushabye’
d. tʃotʃak-tʃotʃak ‘talking without rhyme; tottering’

This fact suggests that the participants are affected by the existing lexical items, especially when they are analogous in shape and/or meaning.[22]

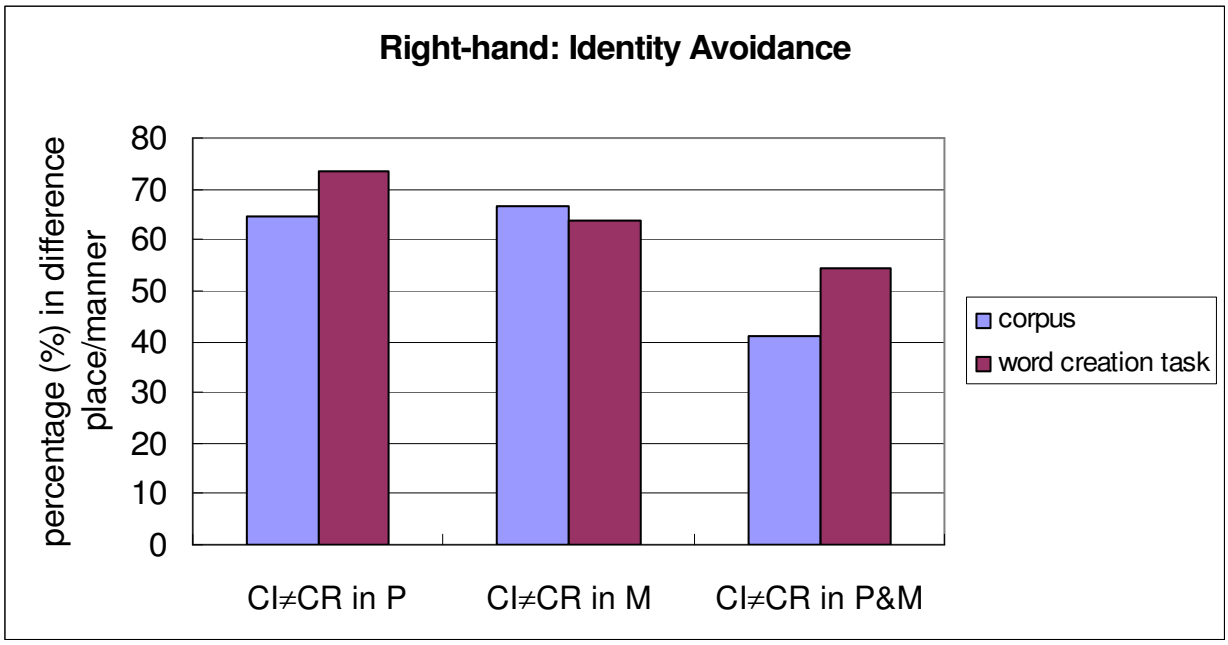
The remaining two types were *apam-**papam*** and *ʌtʃuŋ-**tʃʌtʃuŋ***, both of which have identical CI and CR. However, these account for only 1.48% in the total of 472 tokens.

The experimental results exhibit a greater effect of place than manner in distinguishing the CI and CR. Furthermore, the comparison between corpus and word creation task reveals that the corpus statistic pattern is roughly replicated in the word creation task, while showing the greater effect of place than manner in the word creation task. On the other hand, the Identity Avoidance Effect in the word creation task is even stronger than in the corpus.

[22] Among the other reduplicated forms produced by the participants were six other types which have a corresponding perfect total reduplicated form in the lexicon. However, they come to 9.11% (43 tokens out of 472).

(i)	a.	oloŋ-**tʃ**oloŋ	→	tʃoloŋ-tʃoloŋ	‘in clusters’	14 tokens
	b.	apam-**tʃ**apam	→	tʃapam-tʃapam	‘by driblets’	12 tokens
	c.	otok-**p**otok	→	potok-potok	‘dry and crisp’	6 tokens
	d.	usʌl-**t**usʌl	→	tusʌl-tusʌl	‘talking in a low voice’	5 tokens
	e.	oloŋ-**tʃ^{h}**oloŋ	→	tʃ^{h}oloŋ-tʃ^{h}oloŋ	‘twinkling; wakeful’	4 tokens
	f.	otok-**p’**otok	→	p’otok-p’otok	‘dry and crisp’	2 tokens

(32) Figure 8. CI vs. CR



In sum, the examination of VCVC-bases both in the corpus and the word creation experiment shows that there is a general tendency toward identity avoidance between the CI and its adjacent consonants in terms of place and manner. Between these two factors, place has a greater effect in discriminating the CI and its neighboring consonants, and this appears similar to the restriction in Arabic verbal roots in which a combination of consonants with the same place of articulation are disfavored. The results of the corpus analysis are not simply replicated in the word creation task, but the effect of Identity Avoidance is even stronger in the word creation task.

5. General Discussion

Looking through the corpus reduplication data with an inserted consonant, we found that there is no favored or default consonant inserted in the reduplicant. In line with this finding, we also witnessed that the choice of the inserted consonant depends on the surrounding context. That is,

the CI is chosen with reference to the qualities of the existing base consonants such that the CI tends to differ from the neighboring consonants in terms of the place and manner of articulation.

Due to these findings, I proposed the Identity Avoidance Effect in the choice of CI in reduplication. In order to uphold this proposal, I looked both at the corpus and word creation results, and found that both of them show the Identity Avoidance Effect. The observations hold with both the left-hand and right-hand consonant, called CL and CR, respectively. As was observed, sonorants are generally less preferred in the onset. However, they were still attested in the word creation task by the native speakers when they meet the condition of identity avoidance. Thus we can see that the Identity Avoidance Effect is strong in choosing a CI in reduplication. On the other hand, I do not exclude a possibility that there might be some other factor that is under operation for the choice of CIs. The identity avoidance does not explain everything because we can still see some variation in the results of the corpus and experiment.

The corpus and experiment findings of Korean reduplication are similar to Turkish reduplication data in which the coda consonant of the prefixed CVC syllable tends to differ from the base consonants in terms of segmental resemblance. In both Korean and Turkish reduplication, we find that there is an Identity Avoidance Effect in the sense that the epenthetic consonant in the reduplicant tends to be distinct from the base consonants.

However, it is worthwhile to point out that the choice of epenthetic consonant in Turkish reduplication operates at the segmental level, whereas the choice of inserted consonant in Korean reduplication operates at the featural level. The second C in the CVC prefix of Turkish emphatic reduplication is only taken from the fixed set /p, s, m, r/, whereas the inserted C in the Korean total reduplication is not taken from a fixed set of consonants, rather it is chosen by making reference to the features of the neighboring consonants. On the other hand, the reduplication phenomena from Korean and Turkish are similar in that the inserted consonant is chosen from a

set of consonants in each language, /t, p, tʃ/ in Korean, in the case of the corpus, and /p, s, m, r/ in Turkish. However, in the reduplication experiment with nonce base forms, it was found that Korean speakers have a wider range of choices of epenthetic consonants, i.e. /t, p, tʃ, k, s, m/, whereas Turkish speakers have a more limited choice of epenthetic consonants, i.e. /p, s, m/.

In the meantime, there was no distance effect attested or claimed for the choice of inserted consonants in Turkish reduplication. However, Korean reduplication data seem to show some distance effect. In the preceding sections, I looked both at the left-hand and right-hand context to see whether there is identity avoidance between the inserted consonant and its adjacent segments. I considered VCVC-bases in particular. With VCVC-bases, we obtain a reduplicative form of $\mathbf{VC_2VC_1\text{-}C_iVC_2VC_1}$ in which $\mathbf{C_i}$ represents the inserted consonant (CI), $\mathbf{C_1}$ is a left-hand consonant (CL), and $\mathbf{C_2}$ is a right-hand consonant (CR). We found that there are no cases of CI and CL which share the same place and manner, although we saw an insignificant number of cases of CI and CR which share the same place and manner. This means that CI ($\mathbf{C_i}$) tends to be more distinct from the directly adjacent base consonant ($\mathbf{C_1}$=CL) than from a more distant base consonant ($\mathbf{C_2}$=CR). The consideration of the distance factor puts Korean reduplication facts in parallel with the facts from Arabic verbal roots.

The findings from Korean reduplication support the idea of an Identity Avoidance Effect found in the other languages. In addition, the Korean reduplication analysis argues that the speakers' own grammar does not merely reflect the statistics of the lexicon because the pattern discovered in the lexicon was not just replicated in the experiment to the speakers, but it was also reinforced in the experiment in terms of the identity avoidance and range of possible inserted Cs.

6. Theoretical Implications and Future Directions

The Identity Avoidance Effect suggested for the Korean reduplication in this paper is relevant to the notions of the gradient phonotactics and OCP constraint in the extant theories. Providing evidence from many different languages with respect to the tendency to avoid repetition – local and non-local – Frisch *et al*. (2004) propose a formulation of phonotactic knowledge based on the idea that phonotactic acceptability is a gradient concept which is to be reflected in the pattern of lexical items of a language. In the same vein, Albright (2006) claims that "grammar itself is probabilistic and gradient." (p. 1) Frisch *et al*. point out that the traditional OCP constraint, which they call total OCP, is defective because there are some data it cannot explain. Instead, they suggest the gradient OCP, and this gradient or stochastic notion of OCP is consistent with the findings of this paper.

As for the correlation between the lexicon and the phonotactic constraints, a passage from Frisch *et al*. (2004) is noteworthy: "…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) However, as was seen in the present study, the lexicon cannot be an exact mirror of the speakers' phonotactics.[23] Therefore, whether functional pressures or the emergence of the unmarked motivates the observed identity avoidance effect is one of the controversial questions that remain to be answered.

[23] One of the hypotheses regarding the experiment was that the native speakers of Korean are inclined to choose the same set of consonants for insertion into reduplicant, if they were given the same set of consonants in the base. However, it could not be tested in the current task because among the 2C-base data, only two items for each of the contexts, /l/-/s/ and /ŋ/-/tʃ/, were detected. In addition, for the latter context /ŋ/-/tʃ/ (bases: /atʃaŋ-/ and /ʌtʃuŋ-/), the participants might have been affected by the existing reduplicative forms in the lexicon, /tʃatʃaŋ-tʃatʃaŋ/ and /ʌtʃuŋi-**t'**ʌtʃuŋi/, respectively. For this phenomenon, I can comment that there must be some relevance between the native speakers' lexical knowledge and the phonotactics of the language.

The analysis of the corpus and experimental results in this paper suggest that the native speakers' knowledge does not simply mirror the phonotactics of the lexicon, but it shows a stronger tendency, e.g. in terms of identity avoidance in the total reduplication. Consequently, the evidence in this paper supports the idea that a constraint like the OCP is not categorical but gradient. Further, the effect is not due to the lexical statistics, but due to a universal preference for identity avoidance.

In further research, I plan to apply the feature similarity metric and a related computation test for the corpus and word creation database which I obtained. I would also like to look at the other experiment which was conducted with a multiple-choice task by the same group of participants, the results of which were not included in this paper but are to be analyzed in the future research. In addition, some questions remain: (i) Why does the place factor seem to play a greater role in the dissimilarity between CI and its neighboring consonants?; (ii) Are there similar avoidance patterns found in other cases of C-insertion in Korean?

Appendix

Korean has reduplicative forms as in the following:

(한국어에는 아래 예와 같은 의성어, 의태어들을 포함하는 중첩어가 있습니다.)

Example: (예)

a. 오손-	/oson/	➔	오손도손	/oson-**t**oson/
b. 옹기-	/oŋki/	➔	옹기종기	/oŋki-**tʃ**oŋki/
c. 상글-	/saŋkɨl/	➔	상글방글	/saŋkɨl-**p**aŋkɨl/
d. 생게-	/seŋke/	➔	생게망게	/seŋke-**m**aŋke/

e. 우불- /upul/ ➔ 우불꾸불 /upul-**k'**upul/

f. 알뜰- /alt'ɨl/ ➔ 알뜰살뜰 /alt'ɨl-salt'ɨl/

In the following two sections, we would like to know how you, as a native speaker of Korean, judge the acceptability of novel forms of reduplication, and how you create a reduplicative form by making use of the given portion of word.

(다음에 이어지는 두 종류의 테스트는 여러분이 한국어 화자로서 낯선 중첩형의 적절성을 어떻게 판단하는가와 주어진 단어의 일부를 취하여 새로운 중첩형을 어떻게 만드는가를 살펴보기 위한 것입니다.)

WORD CREATION (낱말 만들기)

Directions: Imagine that the given portion is part of a reduplicated form in Korean. Complete the form by adding another portion that has an inserted (or substituted) segment in the initial position of reduplicant. You are allowed to give more than one acceptable form. (다음은 한국어에 존재할 수 있는 중첩어의 일부입니다. 주어진 형태소에 초성을 삽입하거나, 또는 주어진 형태소에 이미 초성이 존재할 시에는 그 초성을 바꿈으로써 주어진 형태를 완성된 중첩어로 만들어 주십시오. 가능한 중첩어가 두 가지 이상 떠오른다면 그들 모두를 써 주십시오.)

1. ʌntɨl-( )

2. usul-( )

3. patʃik-( )

4. ʌntʃaŋ-( )

5. tuluk-( )

6. otok-( )

7. salk^{h}aŋ-( )

8. otʃak-( )

9. kasam-( )

10. silʌŋ-( )

11. ultʃ’a-( )

12. komtʃik-( )

13. asik-( )

14. atal-( )

15. p’asa-( )

16. utʃʰil-( )

17. ʌŋtʃʰa-( )

18. puŋso-( )

19. atʃ’a-( )

20. ʌul-( )

21. tʃolsa-( )

22. oksu-( )

23. atʃu-( )

24. ʌktʃ’aŋ-( )

25. ʌtʃuŋ-( )

26. kutɬl-( )

27. osam-( )

28. uk’ɨn-( )

29. kaman-( )

30. atʃaŋ-( )

31. usʌl-( )

32. kotaŋ-( )

33. okoŋ-( )

34. tamp^{h}uŋ-( )

35. apam-( )

36. okam-( )

37. motɨn-( )

38. huluk-( )

39. oloŋ-( )

40. ontaŋ-( )

References

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

Albright, Adam. 2006b. Explaining universal tendencies and language particulars in analogical change. Ms. MIT. http://web.mit.edu/albright/www/

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

Bao, Zhiming. 1990. Fanqie Languages and Reduplication. *Linguistic Inquiry* 21-3: 317-350.

Blevins, Juliette. 2004. *Evolutionary Phonology: The Emergence of Sound Patterns*. Cambridge University Press.

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.

Coetzee, Andries and Joe Pater. 2006. Lexically Ranked OCP-Place Construction in Muna. ROA # 842 at www.roa.rutgers.edu.

Coetzee, Andries. 2006. Grammar is Both Categorical and Gradient. ROA # 864 at www.roa.rutgers.edu.

Cutler, Anne, Nuria Sebastián-Gallés, and Brit van Ooijen. 2000. Constraints of Vowels and Consonants on Lexical Selection: Cross-linguistic Comparisons. *Memory and Cognition* 28: 746-755.

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

Frisch, Stefan A. and Bushra Zawaydeh. 2001. The Psychological Reality of OCP-Place in Arabic. *Language* 77: 91-106.

Frisch, Stefan A. 2004. Language Processing and Segmental OCP Effects. In Hayes, Kirchner, and Steriade (eds.), *Phonetically-based phonology*, 346-371. Cambridge University Press. www.cas.usf.edu/~frisch

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.

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

Hay, Jennifer, Janet Pierrehumbert, and Mary Beckman. 2004. Speech Perception, Well-formedness and the Statistics of the Lexicon. In Local, J., R. Ogden and R. Temple, eds., *Phonetic Interpretation: Papers in Laboratory Phonology VI.* Cambridge University Press, pp. 58-74.

Hayes, Bruce. 1986. Inalterability in CV Phonology. *Language* 62: 321-351.

Hayes, Bruce and Zsuzsa Cziraky Londe. 2006. Stochastic Phonological Knowledge: The Case of Hungarian Vowel Harmony. *Phonology* 23: 59-104.

Hume, Elizabeth and Georgios Tserdanelis. 2002. Labial Unmarkedness in Sri Lankan

Portuguese Creole. *Phonology* 19: 441-458.

Hume, Elizabeth and Ilana Bromberg. 2005. Predicting Epenthesis: An Information-theoretic Account. http://www.ling.ohio-state.edu/~ehume/

Jun, Jungho. 2004. Place Assimilation. In Bruce Hayes, Robert Kirchner, and Donca Steriade (eds.), *Phonetically-based Phonology*, pp. 58-86. Cambridge University Press.

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

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

Kim, Soohee. _____. Reduplication and the Mimetic Stratum in Korean. University of Washington. Ms.

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

Kingston, John and Paul de Lacy. 2006. Synchronic Explanation. University of Massachusetts at Amherst and Rutgers University. Ms.

Kirchner, Robert. 2000. Geminate Inalterability and Lenition. *Language* 76: 509-545.

Lombardi, Linda. 2002. Coronal Epenthesis and Markedness. *Phonology* 19: 219-251.

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

McCarthy, John J. 1993. A Case of Surface Constraint Violation. *Canadian Journal of*

Linguistics 38-2: 169-195.

McCarthy, John J 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. pp. 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.

Pulleyblank, Douglas. 1988. Vocalic Underspecification in Yoruba. *Linguistic Inquiry* 19: 233-270.

Selkirk, Elizabeth. 1984. On the Major Class Features and Syllable Theory. In *Language Sound Structures*. Mark Aronoff and Richard T. Oehrle (eds.). 107-136. Cambridge, MA: 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. http://roa.rutgers.edu/files/82-0000/roa-82-yip-3.pdf

Yip, Moira. 1997. Repetition and its Avoidance: The Case of Javanese. http://roa.rutgers.edu/files/83-0000/83-0000-YIP-0-0.PDF