

# Constraint on Proximate Repetition

—With Special Reference to English

Nahm-Sheik Park

## 1. Preliminaries

We sometimes experience the difficulty of telling one twin from the other or of correctly dialing a telephone number containing a long string of identical numerals. Evidently the difficulty here arises from the proximate co-occurrence of similar elements.

A similar phenomenon is observable in human language also with the result that proximate repetition of identical or nearly identical linguistic elements tends to be constrained. In fact, it has been argued elsewhere (Park, 1977a and 1977b) that something like the following principle lies at the root of this constraint on proximate repetition.

The greater the similarity and proximity between elements occurring in an expression, the greater the (processing) difficulty and hence the greater the awkwardness of the expression.

In the remainder of this paper, we will demonstrate that this repetition constraint operates on all levels of linguistic structure, i.e. that proximate repetition of similar elements is maximally constrained on all levels of linguistic structure. Although virtually all of our examples are from English, the underlying assumption throughout the paper is that similar examples abound in all human languages.

## 2. Syntactic and Semantic Data

An abundance of data supportive of our repetition constraint can be adduced from the syntactic and semantic strata of the English language, especially the former. The doubling constraint (Ross) and the like-case constraint (Fillmore) are a mere two examples.

Our constraint does indeed appear to throw considerable light on a fair number of grammatical peculiarities of the English language. As our first example, let us consider the following data.

- (1) a.? The wall will *be being* painted.  
b. The wall will *be getting* painted.
- (2) a.? The wall has *been being* painted.  
b. The wall has *been getting* painted.

The first sentence in each pair here is awkward, apparently on account of the proximate repetition of "be", and is thus likely to be avoided in favor of the second sentence, which does not involve a similar repetition.

The constraint on proximate repetition also appears to be instrumental in explaining a number of rather interesting syntactic neutralizations.

The following data provides us with one such neutralization.

- (3) a. That he is examined here surprises me.  
 b. His being examined here surprises me.  
 (4) a. That he is being examined here surprises me.  
 b. \*His *being being* examined here surprises me.  
 c. His being examined here surprises me.

Since (3 a) transforms into (3 b) via nominalization, we would expect (4 a) to transform into (4 b) by the same process. However, (4 b) involves a contiguous repetition of "being" and is thus very awkward with the result that it is avoided in favor of (4 c), which is (4 b) minus one "being". Thus both (3 a) and (4 a) transform into the identical "His being examined here surprises me".

The following data offers another similar case of syntactic neutralization.

- (5) a. That he sleeps here surprises me.  
 b. His sleeping here surprises me.  
 (6) a. That he is sleeping here surprises me.  
 b. \*His *being sleeping* here surprises me.  
 c. His sleeping here surprises me.

Since (5 a) transforms into (5 b) via nominalization, we would expect (6 a) to similarly transform into (6 b). However, (6 b) involves a proximate repetition of "-ing" and is thus very awkward with the result that it is avoided in favor of (6 c), which is (6 b) minus "being." Thus both (5 a) and (6 a) transform into the identical "His sleeping here surprises me".

Our constraint also provides a principled account for the neutralization exemplified by the following data.

- (7) a. The girl who works here comes from Canada.  
 b. The girl working here comes from Canada.  
 (8) a. The girl who is working here comes from Canada.  
 b. \*The girl *being working* here comes from Canada.  
 c. The girl working here comes from Canada.

Since (7 a) transforms into (7 b) via relative-clause reduction, we would normally expect (8 a) to transform into (8 b) by the same process. However, (8 b) involves a proximate repetition of "-ing" and is thus very awkward so that it is avoided in favor of (8 c). Thus both (7 a) and (8 a) transform into one and the same "The girl working here comes from Canada".

The following data shows that contiguous repetition of prepositions also constitutes a

serious breach of our constraint.

- (9) a. \*He was waiting *for for* her to come home.  
 b. He was waiting for her to come home.

We derive (9 a) by embedding "she comes home" into the slot of "X" in "He was waiting for X". Since the typical filler for "X" is a nominal, "she comes home" is nominalized into "for her to come home" prior to its insertion into this slot. We thus arrive at (9 a), which is extremely awkward on account of the contiguous repetition of "for" and is thus avoided in favor of (9 b), which is (9 a) minus one "for".

The following data is amenable to a similar explanation.

- (10) a. \*He prefers *to leave to to* stay.  
 b. He prefers leaving to staying.

We obtain (10 a) by embedding "he leaves" and "he stays" into the "X" and "Y" slots of "He prefers X to Y". "X" and "Y" being typically nominal slots, the two insert sentences are nominalized into "for him to leave" and "for him to stay" prior to their insertion into their respective slots. Since the subjects of the insert sentences are co-referential with the subject of the matrix sentence, "for him" is erased from both slots. We thus arrive at (10 a), which is extremely awkward on account of the proximate repetition of "to" and is thus avoided in favor of (10 b).<sup>1</sup>

We can also explain why (11 a) below is more awkward than either (11 b) or (11 c).

- (11) a. (?)I was born *in January in* 1940.  
 b. I was born in January, 1940.  
 c. I was born in January of 1940.

The problem with (11 a) is that it involves a proximate repetition of temporal in-phrases. Felt to be awkward on this account, (11 a) is almost always replaced by either (11 b) or (11 c). Note here that (11 b) and (11 c) circumvent the problem of (11 a) by not using the second "in".

It goes without saying that we can similarly account for the following data involving a

<sup>1</sup> Our constraint may help explain why the to-infinitive may not be used as the complement of the preposition "to" in such verb phrases as "look forward to...", "object to...", and "be used to..." while the gerund can be so used. Consider the following sets of sentences.

1. a. I look forward to the wedding.  
 b. I look forward to going to the wedding.  
 c. \*I look forward *to to* go to the wedding.
2. a. I am used to a heavy breakfast.  
 b. I am used to having a heavy breakfast.  
 c. \*I am used *to to* have a heavy breakfast.

In fact, the *to*-infinitive may not be used as the complement of any preposition because such a construction entails an extremely awkward cluster of prepositions.

proximate repetition of locative in-phrases.

- (12) a. (?) He was *in* the kitchen *in* the apartment.  
 b. He was in the kitchen of the apartment.  
 c. He was in the apartment kitchen.

It is interesting at this point to note that, with a pause between the two in-phrases, both (11 a) and (12 a) become far less awkward. But this is not surprising at all since the pause renders the repetition far less proximate and thus practically harmless.

The following data clearly indicates that the repeated prepositions here need not be identical as long as they agree in such features as locativity and temporality.

- (13) a. (?) He was born *on* the 16th *in* January  
 b. He was born on the 16th of January.  
 c. He was born on January 16.  
 (14) a. (?) She lives *at* No. 10 *on* Downing Street.  
 b. She lives at No. 10, Downing Street.

Note that referential specificity plays a role here. The more specific of the two prepositions tends to be retained. The following data may be cited in further support of our claim.

- (15) a. (?) Can you look *at* me *in* the {face/eye}?  
 b. \* Can you look at me  $\phi$  the {face/eye}?  
 c. Can you look  $\phi$  me in the {face/eye}?

Note also that the two phrases here may actually use the same preposition and yet be perfectly natural as long as they do not agree in such features as locativity and temporality. Thus (16) below is perfectly natural although "in" is proximately repeated.<sup>2</sup>

- (16) I was born *in* Seoul *in* 1940.

Let us now turn our attention to the following data.

- (17) a. I bought {all/half/both} of the books.  
 b. I bought {all/half/both}  $\phi$  the books.  
 c. I bought {all/half/both} of them.  
 d. \*I bought {all/half/both}  $\phi$  them.

According to this data, the preposition "of" may be deleted from "all/half/both of X" only when "X" is a bona fide noun (phrase), not when it is a pronoun. This apparently peculiar behavior of "of" is resolved when we realize that "all", "half", and "both" are more pronominal than nominal. For then we can say that the deletion of "of" is blocked only when it would result in a contiguous repetition of pronouns, as in (17 d)

<sup>2</sup> This may mean that meaning carries more weight than form does in the operation of our constraint on proximate repetition.

Our constraint on proximate repetition apparently enables us to better understand why we have conjunction deletion as in the following data.

- (18) a. (?) I met Jim *and* John *and* Bill.  
 b. I met Jim, John and Bill.
- (19) a. (?) Send Jim *or* John *or* Bill.  
 b. Send Jim, John or Bill.

Our explanation here would consist in saying that (18 a) and (19 a) are rather awkward on account of the proximate co-occurrence of coordinate conjunctions so that they tend to be avoided in favor of (18 b) and (19 b) respectively.

Repetition of subordinate conjunctions also tends to render an expression infelicitous. Thus (20 a) below is likely to be avoided in favor of the three alternatives that follow.

- (20) a. (?) I know *that* he thinks *that* it is bad.  
 b. I know that he thinks  $\phi$  it is bad.  
 c. I know  $\phi$  he thinks that it is bad.  
 d. I know  $\phi$  he thinks  $\phi$  it is bad.

Let us now consider the following set of sentences, in each of which the so-called negative element is contiguously repeated.<sup>3</sup>

- (21) a. ?\*Such things are *not not* necessary.  
 b. \*Such things are *unnecessary*.  
 c. Such things are *not unnecessary*.

We may note that the two tokens of the negative element are formally identical in the first two sentences while they are not in the third sentence, which is why the first two sentences are extremely awkward while the third sentence is natural. We may further

<sup>3</sup> Our constraint may also be instrumental in explaining the fact that predominantly relational markers may not be proximately repeated. Let us consider the following data.

1. a. We were aware *of* its existence.  
 b. \*We were aware  $\phi$  its existence.  
 c. \*We were aware *of that* it existed.  
 d. We were aware  $\phi$  *that* it existed.
2. a. I was surprised *at* the news.  
 b. \*I was surprised  $\phi$  the news.  
 c. \*I was surprised *at to* hear the news.  
 d. I was surprised  $\phi$  *to* hear the news.

(1 c) and (2 c) are apparently ungrammatical because both involve a proximate repetition of purely relational markers. On the other hand, (1 d) and (2 d), which are (1 c) and (2 c) minus one of the markers in question, are perfectly grammatical. We may note here that "We were aware *of* the fact *that* it existed" is perfectly grammatical because the two markers in question are no longer so proximately repeated. Note finally that the preference of "Look  $\phi$  *who's* here" over "?Look *at who's* here" is amenable to a similar explanation.

note that the two tokens of the negative element are less proximate in the first sentence than in the second, which is why the first sentence is generally felt to be slightly less awkward than the second.

Our constraint enables us to account for the following data along similar lines also.

- (21) a.\*? I cannot *not* obey him.<sup>4</sup>  
 b. I cannot *disobey* him.  
 c. I cannot *but* obey him.  
 d. I cannot *help* obeying him.

Our constraint also sheds much light on why proximate repetition of derivational suffixes is generally avoided, as is shown quite clearly by the following data.

- (22) a.(?) We examined the problem from the *historical sociological* angle.  
 b. We examined the problem from the angle of historical sociology.  
 (23) a.(?) I am looking for some *diachronic linguistic* evidence.  
 b. I am looking for some evidence from diachronic linguistics.  
 (24) a. \*I called her *hourly daily*.  
 b. I called her hourly every day.  
 (25) a. \*He ran away *cowardlily*.  
 b. He ran away in a cowardly manner.

Note in this connection that an adjective in “-ly” hardly ever allows adverbialization in “-ly”. Thus “\**womanlily*”, “\**soldierlily*”, “\**gentlemanlily*” etc. are simply un-English. Although “*manlily*” is a possible English word according to *Webster's Third International*, it is better avoided if we are not to sound odd or archaic.

It is interesting that such pronominal phrases as “*this one*” and “*that one*” are perfectly natural whereas “*these ones*” and “*those ones*” are slightly awkward. It is also interesting that we frequently replace “*these ones*” and “*those ones*” with “*these*” and “*those*” respectively in order to remove this awkwardness. Note here that, in terms of our constraint, “*these ones*” and “*those ones*” are awkward because of the proximate repetition of the plural suffix “-s” whereas “*these*” and “*those*” as well as “*this one*” and “*that one*” are perfectly natural because they do not involve a similar repetition.

Let us now turn our attention to the following data.

- (26) a. ? For some *reason* or other *reason*, they did not come.  
 b. For some reason or other  $\phi$ , they did not come.  
 (27) a. ? For one *reason* or another *reason*, they did not come.  
 b. For one reason or another  $\phi$ , they did not come.

<sup>4</sup> Incidentally, “I can't *not* obey him” and “It isn't *not* necessary” may be slightly more acceptable than “I cannot *not* obey him” and “It is *not not* necessary” respectively. For “-n't not” involves a less serious violation of our constraint on proximate repetition than does “not not”.

Both (26 a) and (27 a) are very awkward because of the proximate repetition of "reason"; they are thus avoided in favor of (26 b) and (27 b) respectively. Needless to say, the following data can be accounted for in a similar manner.

- (28) a. (?) It was between Sam, on the one *hand*, and Bob and me, on the other *hand*.  
 b. It was between Sam, on the one hand, and Bob and me, on the other  $\phi$ .  
 (29) a. (?) For one *thing*, he is smart; for another *thing*, he is rich.  
 b. For one thing, he is smart; for another  $\phi$ , he is rich.

Similarly explainable are numerous other expressions of enumeration, from which all but the first token of the noun in question may or must be deleted. It is important to note that the distance between the various tokens of the noun in question plays a crucial role here. In general, the following appears to hold. That is, the longer the distance between the tokens of the noun in question, the less likely the deletion. In this respect, the deletion here is very much like anaphoric pronominalization. It is also like anaphoric pronominalization in that all but the first token of the noun in question are affected by the operation.

Note in this connection that our constraint could be used in explaining why parallel structures often steer clear of articles, as is shown in the following data.

- (30) a. (?) The country extends 200 miles from *the* east to *the* west.  
 b. The country extends 200 miles from  $\phi$  east to  $\phi$  west.  
 (31) a. (?) I saw the movie from *the* beginning to *the* end.  
 b. I saw the movie from  $\phi$  beginning to  $\phi$  end.

The awkwardness of both (30 a) and (31 a) is due to the proximate repetition of "the". Note here that not just one token of "the" but both tokens need to be deleted. This is probably because "the" is devoid of any inherent meaning.

It is interesting that we can use our constraint in explaining the distributional differences between "home run" and "homer", as shown in the following data.

- (32) a. He hit another home run.  
 b. He hit another homer.  
 (33) a. ? He hit another *two-run* home *run*.  
 b. He hit another two-run homer.

Note that (33 a) is awkward because "two-run home run" involves a proximate repetition of "run". Thus we can say that "home run" may not occur with impunity in contexts which already have another token of "run", as in (33 a), whereas "homer" is not thus restricted. Explainable along similar lines is the fact that we have "clock-watchers", but not "watch-watchers."

It is also interesting that our constraint makes available a principled explanation as to why tautological expressions are frowned upon and avoided. Note that such tautologies as "\*male prince", "\*female princess", "\*married wife", and "\*unmarried spinster" involve

proximate repetition of identical semantic features in flagrant violation of our constraint.

The following two sentences are also in violation of our constraint.

(34) \*I *gunned* him down with a *gun*.

(35) \*I *penciled* in his name in *pencil*.

Note that these sentences become perfectly natural if we revise their “instrumental” phrases slightly, as in the following sentences.

(34') I *gunned* him down with a *shotgun*.

(35') I *penciled* in his name in *red pencil*.

Revised thus, the repeated elements are no longer identical with the result that the constraint on proximate repetition is not violated in any serious way.

Note in this connection that durational adverbials of the form “all day”, “all week”, etc. are never used with the durational preposition “for”. Thus (36 a) below is very awkward while (36 b) is perfectly natural.

(36) a. \*They drank *for all* night.

b. They drank  $\phi$  all night.

Note here that (36 a) involves a proximate repetition of the durative semantic feature since two neighboring lexical items, i.e. “for” and “all”, denote duration. On account of this proximate repetition, (36 a) is felt to be very awkward and is thus avoided in favor of (36 b), which is (36 a) minus “for”.

Needless to say, we can similarly account for the lack of “for” in such durational expressions as “all the time”, “most of the time”, “much of the time”, and “part of the time”. It is very interesting that this account applies not just to temporal durationals but also to locative durationals. Thus the avoidance of “for” in such locative durationals as “all the way”, “most of the way”, “much of the way”, “part of the way”, and “all the distance” can now be viewed as a consequence of our constraint operating on such expressions.<sup>5</sup>

### 3. Phonological and Orthographical Data

As our first example, let us consider the alternation between the two forms of the indefinite article, i.e. “a” and “an”. We may begin by noting that the two forms are in complementary distribution in that “a” occurs in preconsonantal position while “an” occurs in prevocalic position. Since “a” ends in a vowel while “an” ends in a consonant, this alternation serves to prevent contiguous repetition of either vowels or consonants. It may be in order here to point out that the alternation between “my” and “thy”, on the one

<sup>5</sup> Explainable in similar terms is the fact that adverbials of the type “the whole morning” are seldom used with the durational preposition “for”, as well as the fact that typically durational verbs such as “last” may often be followed by a durational phrase without “for”, as in “The festival lasted (for) three weeks”.



hand, and "*mine*" and "*thine*", on the other, used to serve exactly the same purpose.

It may be noted here that a cluster of either consonants or vowels is on the rare side in human language, especially within a morpheme. Thus a vowel is more likely to be followed by a consonant than by a vowel and vice versa. In our terms, this phenomenon may be a natural consequence of the constraint operating on segmental phonology.

Note in this connection that the constraint on proximate repetition may be responsible for such silent letters as "t" in "castle" "Christmas," "often," "soften", etc. The same thing may be said about the frequent silencing of the italicized letters in such words as "old", "and", "twenty", "clothes", "exactly", and "eighth" /eytθ/. The silent or silenceable letters here are by no means easy to pronounce and get outshouted by their more sonorant neighbors (Hankamer and Aissen, 1974).

Affected in this way are not just consonant clusters but vowel clusters as well. Thus vowels often drop out when immediately followed by another vowel, especially in the process of word formation. In the following data, for example, italicized vocalic segments get deleted in immediate prevocalic position.<sup>6</sup>

- (37) *China* + -ese ⇒ Chinese
- (38) *Genoa* + -ese ⇒ Genoese
- (39) *Mexico* + -an ⇒ Mexican
- (40) *inferno* + -al ⇒ infernal
- (41) *Anti-* + Artic ⇒ Antarctic
- (42) *anti-* + acid ⇒ antacid

We will now consider cases in which proximate repetition is avoided through the insertion of the "neutral" vowel /i/ between two consonants, mostly between a stem-final consonant and a monoconsonantal suffix. Let us first consider the insertion of /i/ between certain verb stems and the past (participial) suffix "-ed". Note that the suffix "-ed" is typically an alveolar stop, i.e. either /t/ or /d/, and that /i/ is inserted only when the verb stem ends in an alveolar stop, i.e. either /t/ or /d/. Thus we can say that /i/ is inserted between two alveolar stops here so as to prevent their contiguous repetition and thus render the sequence pronounceable.

Another case involves the insertion of /i/ between certain stems and the suffix "-s". This suffix, which is typically an apico-alveolar sibilant, i.e. either /s/ or /z/, may be the plural or genitive suffix on nouns or the third-person-singular-present suffix on verbs. Note that /i/ is inserted here only when the stem ends in a sibilant, i.e. /s/, /z/.

<sup>6</sup> Note that "American" comes from "America+an". On the basis of the derivational pattern for the other examples given here, we may claim that, of the two tokens of "a", it is the first one that gets deleted. We may claim with some plausibility that, when we derive "none" from "no"+ "one", it is the first "o" that gets deleted. In support of this claim, we may cite the fact that the vowel in "none" is identical with the vowel in "one", not with the vowel in "no".

/š/, /ž/, /č/, or /j/. Thus we can say that /i/ is inserted between two sibilants so as to prevent their contiguous repetition and thus render the sequence pronounceable.

The tendency to insert /i/ in the final syllables of such words as "castle", "often", "middle", and "bottle" may be explainable in similar terms. In other words, the /i/ insertion here may also serve to break an otherwise awkward consonant cluster. Note here that the frequently observed tendency for foreign learners of English to insert /i/ between the members of a consonant cluster may be interpreted as a consequence of their attempt to break the cluster and render it more pronounceable. Note further that the intervocalic "r" (in British English) as well as the intrusive "r", as in "This pizza (r) is delicious", may be thought of as serving the useful purpose of breaking a vowel cluster.

Let us now consider cases in which proximate repetition is avoided through the dropping or non-use of a sound (sequence), mostly the "-s" suffix. A fairly productive case is provided by the zero genitive, which is used with regular plural nouns, as in "soldiers' wives" and names ending in "s", as in "Socrates' wife". Since the zero genitive is used when the stem-final consonant is "s", we may claim that it serves to help avoid a proximate repetition of "s" sounds and thus render the sequence pronounceable.

The use of the zero-genitive in "for convenience' sake", "for conscience' sake", etc. appears to be similarly motivated, i.e. to avoid the overcrowding of "s" sounds. Note here that a non-zero genitive, i.e. "-s", would be flanked by "s" on both sides.<sup>7</sup>

Another interesting case involving the avoidance of the suffix "-s" is provided by nationality adjectives ending in /s/ or /z/, e.g. "Swiss", "Chinese", etc. We may begin by noting that these adjectives may convert to noun, but that they may not take the plural suffix "-s". In other words, we may have "He is a Swiss" and "He is a Chinese", but not "\*They are Swisses" and "\*They are Chinesees". Note here that "\*Swisses" involve a rather awkward repetition of "s" sounds.

Note in this connection that nationality adjectives ending in sibilants, including /s/ and /z/, allow plural nominalizations of the form "the+nationality adjective", as in "the English" and "the Swiss", but not of the form "the+nationality adjective+-s", as in "\*the Englishes", "\*the Swisses", "\*the Dutches", and "\*the Chinesees". Again this may be motivated to avoid proximate repetition of sibilants, i.e. the repetition of the stem-final sibilant and the suffixal sibilant "-s".

Our constraint also appears to explain why "This is" or "This has" may not contract to "\*This's" while "That is" or "That has" may contract to "That's". Needless to

<sup>7</sup> Note that one of two things occurs here when the "-s" genitive follows a stem-final sibilant. If the stem-final sibilant is an "s" sound, i.e. either /s/ or /z/, the suffix "-s" is apt to zero out. If, on the other hand, the stem-final sibilant is other than an "s" sound, i.e. /š/, /ž/, /č/, or /j/, then the suffix "-s" is retained with the insertion of the neutral vowel /i/ as a buffer between the two sibilants. This is not surprising at all when we remember that our constraint in part says: "The greater the identity between elements in proximate repetition, the more serious the violation of our constraint and hence the greater the need for deletion".

say, the "s" cluster renders "This's" unacceptable.

Haplology affords us still another case in point. It appears that the simplification of "library" and "probably" to "libry" and "proibly" respectively, for example, is motivated to minimize proximate repetition of similar sounds and thus make the words easier to pronounce.

Still another interesting case comes from the silent letter "w" in such words as "write". Note here that /w/ is phonetically similar to /r/, as is attested to by the fact that in children's speech the former often replaces the latter, as in /kway/ for /kray/ "cry". If so, we may say that we have here a contiguous repetition of two similar sounds, i.e. the cluster /wr/, and that we drop /w/ so as to make the pronunciation easier. Thus we may claim that the dropping of "w" here does not differ in nature from the dropping of "b" from "lamb", for example.

The silent letter "w" in such words as "who", "two", etc. may be explainable in the same manner. It may be argued that we have in \*/hwu:/ "who" or \*/twu:/ "two" an awkward cluster, because of the contiguous co-occurrence of /w/ and /u:/, which are phonetically similar to each other, and that we simplify this cluster by dropping /w/.

Note in this connection that foreign learners of English often fail to pronounce /w/ in the sequence /wu(:)/. For example, they often say /ud/ for /wud/ "wood". They also tend to drop /y/ from the sequence /yi (:)/, thus mispronouncing /yi:st/ "yeast" as /i:st/. Such errors as these can now be better understood and explained in terms of our constraint on proximate repetition.

We will now consider an example from the suprasegmental phonology of English. One major characteristic of English suprasegmental phonology is that prominently stressed syllables tend to alternate rather regularly with less prominently stressed syllables. Thus English stress also appears to obey our constraint on proximate repetition.

As an illustration, let us consider the behavior of such monosyllabic prefixes as "mis-", "dis-", and "mal-" with respect to stress assignment. These prefixes tend to be prominently stressed when the immediately following syllable is weakly stressed while they tend to be weakly stressed when the immediately following syllable is prominently stressed. Thus "dis-", for one, is less prominently stressed in immediate pretonic position, as in "disloyal", "dishonorable", "discourteous", etc., than in non-pretonic position, as in "disagree", "disapprove", "disconnect", etc. Note that this serves to rule out contiguous repetition of prominently stressed syllables.

The stress shift observable in the following data is apparently motivated to minimize proximate repetition of prominently stressed syllables.

- (43) a. He is studying Chinese.  
 b. He is studying the Chinese language.
- (44) a. He read Chapter Fourteen.  
 b. He read the fourteenth chapter.

- (45) a. He sleeps in the room dōwnstāirs.  
 b. He sleeps in the dōwnstāirs rōom.

As our final example, let us consider the following data relating to punctuation.

- (46) a. \*My father is a Harvard Ph.D..  
 b. My father is a Harvard Ph.D.  
 c. Is my father a Harvard Ph.D.?  
 d. \*Is my father a Harvard Ph.D.?

Note here that, of the two sentence-final periods in (46a), the first one is for the abbreviation "Ph.D." while the second one is for the declarativeness of the sentence. The contiguous repetition of the two periods here may constitute too serious a breach of our constraint with the result that (46a) is avoided in favor of (46b), which is (46a) minus one of the two final periods.

#### 4. Epilogue

We have demonstrated that proximate repetition of similar elements is maximally constrained in English. It is our position that no linguistic theory would be complete without something like our constraint on proximate repetition.

It is our conviction that our constraint can shed much light on a wide range of linguistic phenomena. For one thing, this constraint affords us a rare insight into the phenomenon of deletion or reduction. Almost all instances of deletion that we can think of serve to reduce or avoid proximate repetition of similar elements. Thus we may claim that the basic motivation behind most types of reduction or deletion is to render the structure in question more observant of our constraint on proximate repetition. Falling under this explanatory umbrella are the instances of deletion involved in pronominalization, reflexivization, equi-NP deletion, conjunction reduction, relative clause reduction, performative deletion; just to name a few.

This paper may have given the impression that the constraint on proximate repetition operates on the intralinguistic level only. However, as suggested by the mention of performative deletion in the preceding paragraph, the constraint does operate on the extralinguistic plane as well. Many interesting extralinguistic examples supportive of our constraint are discussed elsewhere (Park, 1980 and 1981).

We may note at this point that our constraint operates most noticeably at grammatical boundaries such as morpheme boundaries. This may be because a grammatical boundary tends to be highly unstable just as a geologic fault is. In a manner of speaking, a grammatical boundary may be an area prone to linguistic earthquakes.

We may also note that, when an element gets deleted under the influence of the constraint, the deleted element tends to be weaker in meaning or pronunciation than the remaining element(s) with which it is in proximate repetition. This is not surprising in

view of the fact that weakness in meaning and hence in pronunciation is often responsible for reduction or deletion, as in the following data.

(47) Of course⇒'f course⇒'Course

(48) afraid⇒'fraid

This data shows that the preposition "of" and the prefix "a-" may be reduced or deleted altogether. Note here that the noun "course" and the stem "-fraid" cannot be so reduced or deleted because of their heavy meaning content. The weight of meaning and pronunciation as a determinant of reducibility/deletability is discussed in some detail elsewhere (Park, in preparation).

Now let us turn our attention to the fact that one type of proximate repetition is immune to our constraint. When a repetition of elements functions to intensify the meaning of the expression in question, as in reduplications and frequentatives, the repetition does not seem to cause any difficulty at all. Thus the following sentences are perfectly acceptable although they all involve proximate repetition of similar elements.

(49) I was *very, very* happy.

(50) *Verily, verily*, I say unto you...

(51) *Up, up*, and away!

(52) She *laughed and laughed and laughed*.

(53) The story is getting *more and more* interesting.

(54) There are *professors and professors*.

In light of this data, we will have to stipulate that our constraint does not apply when elements are repeated to intensify the meaning expressed. There must be many other ways in which our constraint could be further refined, but its main thrust seems to be basically correct.

Finally, we will cite a few Korean examples in support of our claim that the constraint on proximate repetition is universal. Let us begin by noting that each of the following sentences is awkward because of the proximate repetition indicated by italicization.

(55) a. ?/na-*man* kii-*man* mannassta/ "Only I met *only* him".

b. /na-nin kii-*man* mannassta/ "I met *only* him".

(56) a. (?) /na-*to* kii-*to* mannassta/ "(?)I *too* met him *too*".

b. /na-to kii-*nin* mannassta/ "I *too* met him"

It is interesting that /*pakpaks*a/ "Dr. Pak" sounds very awkward evidently on account of the contiguous repetition of /*pak*/ whereas /*kimpaks*a/ "Dr. Kim" sounds perfectly natural. It may be pointed out in this connection that /*pakpaks*a/ tends to be avoided in favor of /*taktapak*/, which is a Korean transliteration of the English "Dr. Pak". It is also interesting that we derive /*munhaks*a/ "bachelor of arts" by combining /*munhak*/ "arts" with /*haks*a/ "bachelor". Note here that /*munhak*/ + /*haks*a/ = /*munhakhaks*a/, but that

/munhakhaksa/ is avoided in favor of /munhaksa/ because the former involves a proximate repetition of /hak/ and is therefore very awkward<sup>8</sup>.

In contrast, we derive /munhakupaksa/ "Doctor of Literature" straightforwardly by combining /munhak/ "arts/literature" with /paksa/ "doctor". Note that /munhaksa/, but not /munhakupaksa/, involves the deletion of /hak/ and that this apparent peculiarity is readily explainable in terms of our constraint on proximate repetition.

As our last Korean example, let us consider pairs of alternants such as the subject markers /ka/~/ika/, the object markers /lil/~/il/, and the instrumental markers /lo/~/ilo/. Note that, in each pair of alternants here, the first alternant begins with a consonant and occurs in immediate post-vocalic position whereas the second alternant begins with a vowel and occurs in immediate post-consonantal position. Thus the alternation between the two members of each pair here is clearly motivated to help avoid proximate repetition of either vowels or consonants.

Note in this connection that the alternation between the two declarative forms of the copula, i.e. /ta/ and /ita/, is similarly motivated. The alternant /ta/ is used when the copula follows a complement noun ending in a vowel; the other alternant, i.e. /ita/, is used when the copula follows a complement noun ending in a consonant.

## REFERENCES

- Fillmore, Charles J. (1968) 'The Case for Case', in Emmon Bach and Robert T. Harms, ed., *Universals in Linguistic Theory*, 1-88, Holt, Rinehart and Winston, Inc., New York.
- Hankamer, Jorge and Judith Aissen (1974) 'The Sonority Hierarchy,' *Papers from the Parasession on Phonology*, 131-145, Chicago Linguistic Society, Chicago.
- Park, Nahm-Sheik (1977a) *Variables in Sentential Readability with Special Reference to EFL/ESL for Korean Learners*, Georgetown University Ph.D. Dissertation.
- (1977b) 'Structural Density as a Comprehensibility Variable,' *Language Research* 13.2, 95-103, Seoul National University, Seoul. (in Korean)
- (1980) 'The Distance Variable in Linguistic Description—With Special Reference to English,' *Languages and Linguistics* 6, 9-32, Hankuk University of Foreign Studies, Seoul. (in Korean)
- (1981) 'Distance and Length in Language—With Special Reference to English,' unpublished paper read at a linguistics seminar co-sponsored by the East-West Center and the University of Hawaii Department of Linguistics.
- (in preparation) *Weight as a Linguistic Variable*.
- Quirk, Randolph et al. (1972) *A Grammar of Contemporary English*, Longman Group Limited, London.
- Ross, John R. (1972) 'The Doubl-ing Constraint,' *Linguistic Inquiry* 3.1, 61-68.

<sup>8</sup> Similarly /yapyap/ "Miss Yang" is awkward while /kimyap/ "Miss Kim" is not. Note further that /misyap/, a transliteration of "Miss Yang", tends to replace /yapyap/.

Department of English  
College of Humanities  
Seoul National University  
Kwanak-ku, Seoul 151  
Korea

*(Received 31 March, 1982)*