Three Cross-linguistic Tendencies in Reduplication Revised*

Jin-young Tak
Sejong University

Abstract

This paper examines how three languages (i.e., Chichewa, Sesotho, Agta) comply with three cross-linguistic tendencies in reduplication (i.e., Shape Invariance, Unmarkedness, and Identity discussed by Kager (1999)) from the perspective of Typology. Considering that Chichewa is characterized as total (i.e., stem) reduplication, Sesotho as bisyllabic reduplication, and Agta as closed-syllable reduplication, this paper is to shows that there is a conflict among these three universal tendencies in reduplication. For example, Unmarkedness, conspiring a reduplicant to a bisyllabic reduplicative template (Kager 1999), cannot account for the occurrence of stem reduplication. In the same sense, Identity supports the occurrence of stem reduplication. A closed-syllable reduplicative template in Agta violates the Unmarkedness and Identity tendencies. Given these observations, in this paper a language-specific tendency preference mechanism with respect to Shape Invariance, Unmarkedness, and Identity is proposed to account for the occurrence of these three

* I would like to thank to Laura Downing who mailed me several of her papers and helped in the further development of this paper. I developed important ideas on reduplication through those papers (Downing 1997, 1999a, 1999b). All the errors are the author’s responsibility.
types of reduplication. Although the three languages seem to display three different reduplication patterns, three cross-linguistic tendencies in reduplication may be well-kept in these languages by adopting the tendency preference mechanism specific to a language.

Keywords: Shape Invariance, Unmarkedness, Identity, reduplication, reduplicative template, cross-linguistic tendency, bisyllabic, stem, closed syllable, tendency preference mechanism

1. Introduction

This paper investigates reduplication of three languages – Chichewa, Sesotho, and Agta – considering three cross-linguistic tendencies in reduplication, Shape Invariance, Unmarkedness, and Identity proposed by Kager (1999).\(^1\)

In general, reduplication patterns are grouped into two categories: total (i.e., stem) reduplication and partial reduplication. For instance, Chichewa displays total reduplication while Sesotho and Agta are characterized as partial reduplication. However, Sesotho and Agta are different in such a way that Sesotho is characterized as a bisyllabic template for reduplication (i.e., partial reduplication), while Agta as closed-syllable reduplication (i.e., partial reduplication).

---

\(^1\) Chichewa and Sesotho are the Bantu languages, spoken in the southern half of Africa (Bleek 1962, Greenberg 1963, Guthrie 1967-71). Bantu is a group of African languages forming a subdivision of the Benue-Niger division of the Niger-Congo branch of the Niger-Kordofanian language family (Greenberg 1963, Guthrie 1967-71, Williamson 1989). In general, Bantu languages are indicated by a geographical zone, labelled A, B, C, D, E, F, G, H, K, L, M, N, P, R, and S. Furthermore, they are divided into language groups, numbered 10, 20, 30, 40, and so on, and then each group has up to nine languages, numbered 1-9 (Wald 1987, Nurse 2002). In this classification, Chichewa and Sesotho are classified as N31 and S33, respectively. Agta is a language used in the Philippines.
As mentioned by many phonologists, reduplicants not only tend to have unmarked structures despite the general phonotactic possibilities of the language (Steriade 1988, Kager 1999), but also tend to preserve phonological identity with the base (Wilbur 1973, Kager 1999). Moreover, Kager (1999) proposes that they have a tendency not to copy the prosodic unit in the base. Instead, there is a reduplicative template that regulates the size of reduplicants. However, this paper illustrates that these three cross-linguistic tendencies in reduplication such as Shape Invariance, Unmarkedness, and Identity fail to account for three different types of reduplication (i.e., bisyllabic reduplication, closed-syllable reduplication, and stem reduplication) simultaneously. Therefore, this paper suggests that in Chichewa Identity is imposed on Unmarkedness, favoring the full copy of a base. As for Sesotho, Unmarkedness is preferred to Identity to derive bisyllabic reduplication. Agta with closed-syllable reduplication, Shape Invariance overrides the other tendencies since a closed syllable is a marked structure and also part of the base (not the whole base).

This paper is organized as follows: Section 2 presents the data on reduplication for three languages: Chichewa, Sesotho, and Agta. Section 3 introduces three cross-linguistic tendencies in reduplication; the new approach corporated with three cross-linguistic tendencies is proposed by means of a tendency preference mechanism. Section 4 offers a conclusion and summary of this paper.

2. Data

2.1. Chichewa Data

Chichewa (N 31) is a Bantu language, spoken in Malawi and neighboring countries. Its native speakers are up to over four millions (Williamson 1989, Nurse 2002).
Consider the following data (Myers-Scotton 1980a, 1980b; Mtenje 1988; Rogers 1989; Kulemeka 1993; Tak 2003, 2005).2

(1) Stem Reduplicated Form Gloss

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a. dyap</td>
<td>idya+idya</td>
<td>‘eat’</td>
</tr>
<tr>
<td>mwap</td>
<td>imwa+imwa</td>
<td>‘drink’</td>
</tr>
<tr>
<td>swap</td>
<td>iswa+iswa</td>
<td>‘break’</td>
</tr>
<tr>
<td>b. lembap</td>
<td>lemba+lemba</td>
<td>‘write’</td>
</tr>
<tr>
<td>gulap</td>
<td>gula+gula</td>
<td>‘buy’</td>
</tr>
<tr>
<td>yanap</td>
<td>yan+yanap</td>
<td>‘see’</td>
</tr>
<tr>
<td>c. lembelap</td>
<td>lembela+lembela</td>
<td>‘write to/for’</td>
</tr>
<tr>
<td>gulidwap</td>
<td>gulidwa+gulidwapa</td>
<td>‘be bought’</td>
</tr>
<tr>
<td>yanapan</td>
<td>yanana+yanapan</td>
<td>‘look’</td>
</tr>
<tr>
<td>yanatanista</td>
<td>yanatanista+yanatanista</td>
<td>‘cause to look’</td>
</tr>
</tbody>
</table>

The data in (1a) are monosyllabic words, while the ones in (1b) are bisyllabic words. The ones in (1c) are words composed of more than three syllables. The reduplication patterns illustrated by (1) clearly show that the reduplication patterns in Chichewa are of total stem reduplication. In Bantu, the constraint that all the words are realized as bimoraic, monosyllabic words cannot be surfaces as they are. As a result, they are augmented by [i] in Chichewa.3 Then, they are reduplicated. It means that all elements of verb stems are reduplicated.

2.2. Sesotho Data

Southern Sesotho (S33) is a Bantu language spoken in the Kingdom of Swaziland in southern Africa (Williamson 1989, Nurse

---

2 In Chichewa, ‘repetitivety’ is conveyed via reduplication.
3 This is further discussed in 3.2. Unmarkedness.
The reduplication data in (2) are based on Doke (1957), Guma (1971), McNally (1990), and Tak (2003:271).4

(2)  

<table>
<thead>
<tr>
<th>Stem</th>
<th>Reduplicated Form</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>ja</td>
<td>jae+ja</td>
</tr>
<tr>
<td>fa</td>
<td>fae+fa</td>
<td>‘give’</td>
</tr>
<tr>
<td>kho</td>
<td>khoe+kho</td>
<td>‘pluck’</td>
</tr>
<tr>
<td>b. aba</td>
<td>aba+aba</td>
<td>‘divide’</td>
</tr>
<tr>
<td>ila</td>
<td>ila+ila</td>
<td>‘abstain’</td>
</tr>
<tr>
<td>esta</td>
<td>esta+esta</td>
<td>‘do’</td>
</tr>
<tr>
<td>fahla</td>
<td>fahla+fahla</td>
<td>‘sprinkle’</td>
</tr>
<tr>
<td>c.</td>
<td>robala</td>
<td>roba+robala</td>
</tr>
<tr>
<td>qutela</td>
<td>qute+qutela</td>
<td>‘finish’</td>
</tr>
<tr>
<td>tsamaea</td>
<td>tsama+tsamaea</td>
<td>‘walk’</td>
</tr>
</tbody>
</table>

The monosyllabic verb stems in (2a) are reduplicated with an infixing vowel [e] in order to satisfy the bisyllabic template in reduplication. The bisyllabic verb stems in (2b) are totally reduplicated; with the polysyllabic verbs having more than two syllables, only the first two syllables are reduplicated, as shown in (2c). From the observations of the above data, it might be defined that unlike Chichewa Sesotho reduplication is characterized as partial reduplication, i.e., bisyllabic reduplication.

2.3. Agta Data

Reduplication patterns in Agta, a language spoken in Philippines, copies segments that stand in closed syllables (Marantz 1982, Kager 1999). This is illustrated in (3).

---

4 In this language, a verb is reduplicated to convey the meaning of ‘a small scale’.
(3) Stem | Reduplicate Form | Gloss
---|---|---
 a. bari | bar-bari | ‘body’
b. wakay | wak-wakay | ‘thing lost’
c. takki | tak-takki | ‘leg’
d. ulu | ul-ulu | ‘head’

In this language, reduplicants cannot be defined as a string of segments such as a consonant plus a vowel (CV or CVV). Based on the data in (3a-c), it may be proposed that in Agta the reduplicants are the first three segments, i.e., CVC. However, this assumption fails to account for the data in (3d). For that reason, instead of mentioning a string of segments, a prosodic reduplicative template such as a closed syllable should be adopted to account for the reduplication patterns in Agta. In other words, the reduplicant in (3d) is a closed syllable regardless of the syllable structure in the base since the reduplicant consists of the segment /l/ that is copied from the base’s onset of the second syllable.

As seen in the data from (1)-(3), the three languages in this paper can be categorized as two subgroups with respect to their reduplication patterns: Chichwa is categorized as a language displaying total (stem) reduplication, and Sesotho and Agta as partial reduplication. Then, partial reduplication in Seotho and Agta is further divided into a language with bisyllabic reduplication and a language with closed-syllable reduplication, respectively.

3. Analysis

Kager (1999) proposes three cross-linguistic tendencies in reduplication: Shape Invariance, Unmarkedness, and Identity. In this

---

5 In Agta, reduplication encodes the meaning of ‘plurality’.
section, these three tendencies are introduced with his terms and how they work for stem reduplication, bisyllabic reduplication, and closed-syllable reduplication.

3.1. Shape Invariance

Shape Invariance, one of the core notions in reduplication, is referred to the fact that there is no one-to-one correspondence with prosodic units in the base. It is defined as follows:

(4) Shape Invariance (Kager 1999: 199)
Reduplication tends to be defined in prosodic units independent of the base. That is, reduplicants do not copy prosodic constituency from the base.

Given this, reduplicants do not simply copy the string of segments or constituents in the base. Instead, reduplicants are related to Shape Invariance involving a notion of a reduplicative template and controlling the size of a reduplicant by facilitating prosodic units such as a syllable or a foot.

In order to find out how Shape Invariance works, examine reduplication patterns in the following two languages, Nookta, a language spoken in British Columbia around and on Vancouver Island, and Diary, an extinct Australian Aboriginal language of South Australia from McCarthy and Prince (1986, 1995) and Kager (1999: 195-196).

(5) Nootka\(^6\)

<table>
<thead>
<tr>
<th>Stem</th>
<th>Reduplicated</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>čims-ʔiːh</td>
<td>či-čims-ʔiːh</td>
<td>‘hunting bear’</td>
</tr>
</tbody>
</table>

\(^6\) Reduplication in Nootka and Diyari means ‘plurality’ of a noun.
As seen in (5) and (6), Nootka reduplicants are characterized as an open syllable, while Diyari reduplicants as a binary foot. In details, even though Nootka allows a closed syllable (i.e., čims-) as in the base čims-ʔi:h, this language permits only an open syllable reduplicant (i.e., či). Therefore, in Nootka the reduplicative template is an open syllable; the reduplicant in (5a) is surfaced as part of the first syllable in the base, či. Different from Nootka, Diyari undergoes binary foot reduplication. In this language, reduplication does not simply copy the first two syllables of the base. The second syllable of the reduplication should be open (i.e., -pa-) even though the one of the base is closed (i.e., -par-). The type of a prosodic unit in Diyari is referred to a foot. Therefore, the Nootka and Diyari data show how different templates (i.e., shape invariants) are adopted in different languages.

The constraint Shape Invariance is also imposed on other languages. As for Chichewa where the reduplicants are the base, the reduplicative template is a whole stem; Sesotho, characterized as having partial reduplication, has a bisyllabic reduplicative template. This is seen in (7).

(7) A Reduplicative Template in Bantu Languages

<table>
<thead>
<tr>
<th>Template</th>
<th>Chichewa</th>
<th>Sesotho</th>
<th>Agta</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stem</td>
<td>Closed</td>
<td>Bisyllabic</td>
<td>Closed Syllabic</td>
</tr>
</tbody>
</table>
3.2. Unmarkedness

Reduplicants also tend to have unmarked structures, as compared to the phonotactic constraints specified in a language. In terms of Unmarkedness, a bisyllabic structure is less marked than other syllable structures (i.e., monosyllabic or trisyllabic). This is well-documented in much previous research. The term ‘minimal word’ (McCarthy & Prince 1986, McNally 1990, Kenstowicz 1997) shows how Unmarkedness is implemented in languages. ‘Minimal word’ is a phonological constraint that the surface realization should be at least two syllables long. In the rare cases of verb roots that comprise only one syllable, including an inflectional final suffix -a, in Bantu a repair strategy is undergone in order to obey ‘minimal word’ (Tak 2005). The surface realization of the monosyllabic words from four different Bantu languages is illustrated in (8) (Myers 1980a, McNally 1990, Park 1997, Hyman, Inkelas & Sibanda 1998, Tak 2005: 373).

<table>
<thead>
<tr>
<th>Language</th>
<th>Underlying Form</th>
<th>Imperative Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Chichewa</td>
<td>/dya/</td>
<td>[idya]</td>
</tr>
<tr>
<td>b. Kihehe</td>
<td>/la/</td>
<td>[lyaa]</td>
</tr>
<tr>
<td>c. Ndebele</td>
<td>/dla/</td>
<td>[yi-dla]</td>
</tr>
<tr>
<td>d. Sesotho</td>
<td>/ja/</td>
<td>[e-ja]</td>
</tr>
</tbody>
</table>

Chichewa, Ndebele, and Sesotho epenthesize [i], [yi], and [e], respectively, to derive a bisyllabic surface form. In contrast, Kihehe lengthens the final inflectional suffix -a since a surface verb stem cannot be monomoraic because of the minimal word constraint.

---

7 Kihehe (G60) is a Bantu language of the Bena-Kinga group, spoken in Tanzania; Ndebele (S40) is a language in the Nguni group of Bantu that is spoken in Zimbabwe (Wald 1987, Nurse 2002).
These processes are applicable to stems having only one syllable, not to stems having two or more syllables. The main reason why they occur depends on the fact that these epenthetic segments are attached directly to the verb stem, in obedience to a general surface word restriction, ‘minimal word’. As described in constraint ‘minimal word’, Unmarkedness is very powerful in phonology.

This Unmarkedness tendency also plays an important role in reduplication phenomena. The definition of Unmarkedness is introduced in (9).

(9) Unmarkedness (Kager 1999: 199)
Reduplicants tend to have phonologically unmarked structures vis-à-vis the phonotactics of the language. That is, reduplicants display a subset of the phonotactic options generally allowed in the language.

In Sesotho, the unmarked structure, a bisyllabic foot, emerges in the reduplicant. This clearly shows their preference for tendency Unmarkedness.

3.3. Identity

Identity is explicitly described in (10), following Kager (1999: 199).

(10) Identity (Kager 1999: 199)
Reduplication tends to preserve full phonological identity with the base. That is, reduplication may involve overapplication and underapplication of phonology to preserve such identity.

Javanese, a language spoken in the central and eastern part of the
island of Java in Indonesia), shows how Identity takes a role in reduplication (Kager 1999: 198, Inkelas & Zoll 2005:137-138).8

(11) Underlying | Surface | Reduplicated | Gloss
----------------|---------|--------------|---------
a. donga | dongə | dongə-dongə | ‘player’
|         | * donga-dongə |
b. donga-ne | donga-ne | donga-donga-ne | ‘my player’
c. meja | mejə | mejə-mejə | ‘table’
|         | * mejə-mejə |
d. meja-mu | meja-mu | meja-meja-mu | ‘your table’

What is interesting in (11) is that Javanese the /a/~/ə/ alternation affects reduplication patterns with respect to Identity. In Javanese, a word-final /a/ raised into /ə/. According to this phonological rule, /a/ surfaces as [a] as it is before suffixes as shown in (11b) and (11d); however, it rises to /ə/ word-finally as in (11a) and (11c). The reason why in (11a) and (11c) the reduplicant does not undergo the /a/~/ə/ alternation while in (11a) and (11c) the reduplicant does is based on the observation that the reduplication patterns in Javanese has a strong tendency to preserve Identity with the base. Therefore, the forms *donga-dongə and *meja-mejə disobey Identity between the base and the reduplicant; their surface forms are *dongə-dongə and *mejə-mejə, respectively.

Again, the Identity tendency, reduplicants tend to preserve phonological identity with the base (Kager 1999: 198). Along with this line, as for stem reduplication in Chichewa, segmental and prosodic identity of the reduplicant and the base is obvious, while in the case of partial reduplication in Sesotho and Agta tendency Identity of the reduplicant is faded away. Chichewa satisfies tendency Identity but disobey Unmarkedness since their reduplicant

8 Javanese reduplication refers to ‘plurality’.
is characterized as a whole stem. However, in Sesotho tendency Markedness dominates tendency Identity, bisyllabic reduplication occurs. Then, Agta has a different tendency preference in order to derive closed-syllable reduplication. In this language, Shape Invariance is the one which is the most preferred tendency. This is discussed in more detail in the next subsection.

3.4. Three Cross-linguistic Tendencies in Reduplication

In subsections 3.1.-3.3., the three cross-linguistic tendencies (i.e., Shape Invariance, Unmarkedness, and Identity) were introduced. However, considering the data from Chichewa, Sesotho, and Agta, it becomes clearer that there exists a conflict among these three tendencies.

Shape Invariance is the tendency which always takes an important role in the reduplication processes since all the reduplicants should be defined in terms of prosodic units. According to Unmarkedness which ensures the occurrence of prosodic unmarked structure, a closed syllable is more marked syllable than an open syllable; a bisyllabic structure is unmarked than any other syllable structures. Therefore, Unmarkedness mitigates stem reduplication and closed-syllable reduplication; the optimal reduplicants within Unmarkedness are composed of two open syllables. In contrast, Identity which prefers the preservation of the base in reduplicants is more likely to produce stem reduplication instead of bisyllabic reduplication.

As for Chichewa where reduplicants are the base, Identity is preferred to Unmarkedness. Of course, Shape Invariance overrides Unmarkedness. In contrast, in Sesotho in order that reduplicants are never longer and shorter than two syllables no matter how long the base is, tendency Unmarkedness and Shape Invariance should be favored over Identity.

Then, how can we account for Agta whose reduplication is
characterized as a close-syllable reduplicative template? Additionally, as shown in (3), to convey ‘plurality’ Agta undergoes close-syllable reduplication. In the example bari ‘body’, the first three segments (i.e., CVC) are copied, resulting in barbari ‘bodies’. However, the word beginning with a vowel instead of a consonant, the segments which are reduplicated are not the first three segments, but the first two segments (i.e., VC). This is seen in the derivation of ululu ‘heads’ from ulu ‘head’. Based on these examples, it is clearly stated that reduplication in Agta is not a simple constituent copying, but involves a prosodic shape invariant in the reduplicants, a closed syllable (Kager 1999: 196). Therefore, the present study proposes that in this language Identity is not dominant since the reduplicants are monosyllabic, not a whole stem. Furthermore, Unmarkedness is also less preferred; the reduplicant in this language is neither an open syllable nor a bisyllabic structure. Therefore, Shape Invariance is favored over the other two tendencies.

Considering the observations above, this paper proposes that the three universal tendencies do not exist in the same tier. Rather, they exists freely, ready to move, as shown in (12).

(12) a. Overall Three Cross-linguistic Tendencies in Reduplication

Among these three cross-linguistic tendencies in reduplication, Shape Invariance is the one that is always highly ranked in the tendency preference mechanism; the reduplicant should be described in terms of a template, which requires the notion of a
prosodic shape invariant (Kager 1999). Additionally, Identity and Unmarkedness cannot stand highly in the same tier because of their own nature. For these reasons, we find only three logically possible rankings as in (13):

(13) Factorial Typology of Tendencies
   a. Unmarkedness, Shape Invariance → Identity
   b. Identity, Shape Invariance → Unmarkedness
   c. Shape Invariance → Unmarkedness, Identity

All languages fall into one of three types of typology of the cross-linguistic tendencies in reduplication.

Then, how does a specific language get the relevant tendency preference mechanism? If there is a conflict or an inconsistency among the three tendencies because of the phonotactic or morphological constraints in a specific language, the tendencies move and are allocated in the right position in the preference mechanism in order to fulfill the constraints.

For example, Chichewa with stem reduplication is analyzed as having Identity override Unmarkedness. Therefore, in this language the Shape Invariance and Identity tendencies move up and apply first; Unmarkedness runs on empty, taking no role. This is illustrated in (14).

(14) The Language-specific Tendency Preference Mechanism for Chichewa

\[ \text{Identity, Shape Invariance} \rightarrow \text{Unmarkedness} \]

As seen in (14), in Chichewa in order to get stem reduplication, Identity and Shape Invariance dynamically move upward, standing in a dominant position above Unmarkedness.
Now consider Sesotho cases, which display bisyllabic reduplication. Different from Chichewa, the tendencies which move are Unmarkedness and Shape Invariance; they override Identity. This is visualized in (15).

(15) The Language-specific Tendency Preference Mechanism for Sesotho

Unmarkedness, Shape Invariance

Identity

Since in Sesotho Unmarkedness and Shape Invariance is favored over Identity, the unmarked structure, i.e., a bisyllabic reduplicative template, is derived.

The next language Agta which has a closed-syllable reduplicative template shows the different preference mechanism from the one of Chichewa with stem reduplication and Sesotho with bisyllabic reduplication. This is shown in (16).

(16) The Language-specific Tendency Preference Mechanism for Agta

Shape Invariance,

Unmarkedness, Identity

In Agta Shape Invariance is more preferred than Unmarkedness and Identity; a reduplicant in this language is neither bisyllabic nor an open syllable. The only tendency considered in this language is Shape Invariance, which governs the prosodic size of a reduplicant.
4. Conclusion

This paper starts the assumption that if a reduplicative template is a marked structure (i.e., a heavy syllable), it increases Markedness in the reduplicant as compared to the base. Kager (1999: 200) proposes that it does not increase markedness since the proper basis for comparison is not the actual base, but rather than the overall prosodic possibilities of the language. For example, the marked CVC reduplicant in Agta is still considered as not violating the overall phonology of the languages; there is no inconsistency between Shape Invariance and Unmarkedness. However, it may be admitted that there might be still favor the unmarked closed syllable. Therefore, this paper proposes the more elaborated definitions. The present study suggests that three cross-linguistic tendencies act together as a group in the process of reduplication; a language facilitates its own mechanism to get the right tendency preference mechanism to fit into its phonotactic or morphological constraints. For example, any language whose reduplication patterns are characterized as stem (total) reduplication like Chichewa, Identity starts to move and overrides Unmarkedness. In a language that has bisyllabic reduplication like Sesotho, Unmarkedness is preferred to Identity. Furthermore, a language like Agta whose reduplication is categorized as closed-syllable reduplication runs the tendency preference mechanism such as Shape Invariance >> Unmarkedness, Identity.

By adopting the tendency preference mechanism specific to a language, this paper comprehensively accounts for the different reduplication patterns in different languages.
References


Kulemeka, A. 1993 *Chichewa I.* Bloomington, IN: Indiana University.


