LANGUAGE AND LINGUISTICS 13.6:1051-1093, 2012
2012-0-013-006-000312-1

Variations in Kavalan Reduplication*

Hui-shan Lin
National Taiwan Normal University

This paper provides an analysis of the reduplicative patterns of Kavalan, an endangered Formosan plains tribe language spoken by fewer than one hundred people on the eastern coast of Taiwan. Kavalan reduplication is special in that the reduplicant takes several distinct shapes depending on the initial syllable of the base (Lee 2009). Within the framework of Optimality Theory (Prince & Smolensky 1993[2004], McCarthy & Prince 1993), this paper shows that Kavalan reduplication is torn between copying a prosodic unit from the base and maintaining an invariant shape. An analysis based on Coetzee's (2006) OT variation model is proposed to account for the variations and predict the relative frequency of the variants.

Key words: Kavalan, reduplication, variation, prosodic faithfulness, Optimality Theory

1. Introduction

As in many Formosan languages, Kavalan, an endangered Formosan plains tribe language spoken by fewer than one hundred people on the eastern coast of Taiwan, displays rich varieties of reduplication in word formation. Kavalan reduplication is mentioned in Li (1982, 1996), J. Lin (1996), Chang (2000), Li & Tsuchida (2006), Lee (2007, 2009, 2010), and examined in depth in the latter two works. According to Lee (2009), except for lexicalized reduplication, whose base is no longer in the synchronic morphology (e.g. *ßayßay* 'a type of small bee', **ßay*) and C*a*- reduplication,[1] formed straightforwardly by the copy of the onset of the initial consonant of the base followed by the fixed vowel *a* (e.g. <u>*sa*</u>~*suani* 'the youngest sibling' < *suani* 'younger siblings'), Kavalan has two other types of reduplication termed monosyllabic and disyllabic reduplication. The reduplicants of monosyllabic reduplication are of the shape CV-, VC-,

* I would like to express my gratitude to my Kavalan consultant Sameg Engi (林阿份) for carefully going through the data with me. I also thank Fuhui Hsieh and Joy J. Wu for their respective help. Finally, I thank the two anonymous reviewers whose valuable comments have helped improve the content of this paper greatly. All possible errors are my own responsibility.

[1] The term C*a*- reduplication is coined by Blust (1998).

CCV-, and CVC- and those of disyllabic reduplication are of the shape CVV-, CVCV- and VCV-, as illustrated by data from Lee (2009) in (1). (In this paper, the reduplicants are underlined and are separated from the base by ~, other morpheme boundaries are marked by -.)

(1) *Examples of monosyllabic and disyllabic reduplication*

a. Monosyllabic reduplication

i.	CV-	nanum	'to drink'	>	m-na~nanum	'to keep drinking'
ii.	VC-	iŋtuʁ	'burned (from overcooking)' >		su-iŋ~iŋtuʁ ʔmaj-su	'Your rice has a burned smell.'
iii.	CCV-	ɬtiq	'to jump'	>	mu-ɬti~ɬtiq	'to keep jumping'
iv.	CVC-	tuŋuz	'to bark'	>	m-ʁi-tuŋ~tuŋuz	'to keep barking'

b. Disyllabic reduplication

i.	CVV-	βaut	'fish'	>	m-ʁi-βau~βaut	'to keep fishing'
ii.	CVCV-	βuɾan	'moon, month'	>	pi-βuɾa~βuɾan	'every month'
iii.	VCV-	m-uʁiŋ	'to weep'	>	m-uʁi~uʁiŋ	'to keep weeping'

Semantics plays no role in determining the choice of different reduplicant shapes because Kavalan reduplication exhibits the so called "*pattern conflation*" in which bases can have attached different reduplicative affixes without changing the meaning (Lee 2007, 2009).[2] The lack of the role of semantics is also obvious by examining (1) because the function of continuity can be denoted by different types of reduplicants such as CV- (1ai), CCV- (1aiii), CVC- (1aiv), CVV- (1bi), and VCV- (1biii).

Lee (2009) proposes, instead, that the choice of the various reduplication patterns is determined by the first syllable of the base. Lee remarks that formation of reduplication in vowel initial bases is rather straightforward (cf. (2a)): bases that begin with a VCV sequence (henceforth, V.CV-BASE) would yield a VCV reduplicant while those with the first syllable as VC (henceforth, VC-BASE) will yield a VC reduplicant. Consonant initial bases, on the other hand, exhibit variations (cf. (2b)). Bases with the first syllable as CVC or starts with a CCV sequence (henceforth, CVC-BASE and CCV-BASE), will yield reduplicants of CVC- and CCV-, respectively.[3] For bases with the first syllable as CV

[2] Please refer to Lee (2009) for a detailed discussion of the semantic functions of Kavalan reduplication.

[3] Lee (2009:132) mentions that "a trisyllabic word may yield a CV.VC reduplicant shape if the second syllable is closed." The only example provided in Lee (2009) that involves CV.VC reduplication is the southern Min loanword *kiamsay* "salted pickle" (*kiâm-chhài*). As no other examples of CV.VC reduplication could be found in the literature, we would like to consider the example to be an exception.

(also referred to as CV-BASE), if the second syllable of the base is onsetful (henceforth, CV.C-BASE), the reduplicant can be CV-, CVC- or CVCV-, though the last pattern is less common; if the second syllable of the base is onsetless (henceforth, CV.V-BASE), the reduplicant can only be CVV-.

(2) *Correlation between base type and reduplicant shape*

a. Vowel Initial Base:

i. V.CV-BASE → V.CV$_{\mathrm{RED}}$ (e.g. uman 'again' > m-<u>uma</u>~uman 'again and again')

ii. VC-BASE → VC$_{\mathrm{RED}}$ (e.g. iŋtuʁ 'burned (from overcooking)' > su-<u>iŋ</u>~iŋtuʁ ʔmaj-su 'Your rice has a burned smell.')

b. Consonant Initial Base:

i. CVC-BASE → CVC$_{\mathrm{RED}}$ (e.g. mazmun 'many$_{[+\mathrm{human}]}$' > sia-<u>maz</u>~mazmun 'getting more and more')

ii. CCV-BASE → CCV$_{\mathrm{RED}}$ (e.g. ktun 'to cut off' > ma-<u>ktu</u>~ktun 'to keep cutting')

iii. CV-BASE:

α. CV.C-BASE → CV$_{\mathrm{RED}}$ (e.g. ma-tamaz 'sharp' > ma-<u>ta</u>~tamaz 'very sharp')
→ CVC$_{\mathrm{RED}}$ (e.g. ʁamaz 'to cook' > m-<u>ʁam</u>~ʁamaz 'to keep cooking')
→ CV.CV$_{\mathrm{RED}}$ (e.g. tuɾis 'spotty' > sa-<u>tuɾi</u>~tuɾis ja quɬus-su 'Your dress is colorful.')

β. CV.V-BASE → CV.V$_{\mathrm{RED}}$ (e.g. muaza 'many$_{[-\mathrm{human}]}$' > sia-<u>mua</u>~muaza taquq 'getting more and more chickens')

Lee's observation as summarized in (2) is insightful, but exactly why a base initial syllable would influence choice of reduplicant shape remains unexplained. Issues like why the reduplicant varies in size (between monosyllabic and disyllabic) and why CV.C-BASE would yield various reduplicant patterns (but a similar base, CVC-BASE, does not) also deserve investigation. Besides, the influence of the base initial syllable on the selection of reduplicant shape also seems to imply that Kavalan reduplication lacks the universal property of *shape invariance* (Moravcsik 1978), which refers to the tendency for the reduplicant to have an invariant shape that has no one-to-one relation with a prosodic constituent in the base. What role *shape invariance* plays in Kavalan also requires further examination.

Within an OT framework, this paper re-examines monosyllabic and disyllabic reduplication in Kavalan (collectively referred to as Kavalan reduplication hereafter). This paper argues that the default reduplicant size is bimoraic. It is shown that though the reduplicant inclines to maintain an invariant size, it also tries to copy the base initial syllable, which varies in shape. The result of the competition between the two contradictory forces is that some bases yield various reduplicant shapes but the others do not. The prosodic faithfulness constraint (McCarthy 2000) is shown to play an important role in conditioning the occurrence of variation because an undersized (monomoraic) reduplicant can only surface if it can help improve the IR matching of syllable structure. An analysis based on the ROE approach (Coetzee 2006) is posed to account for the variations in Kavalan reduplication.

This paper also examines and presents examples not previously considered in Lee (2009) (including second-hand data cited from *Kavalan Dictionary* compiled by Li & Tsuchida 2006 and from Lee 2007, Lee 2010, J. Lin 1996, Chang 2000, Shen 2005, Jiang 2006, D. Lin 2006, and Hsieh 2007[4] as well as some first hand data collected by myself with a native speaker of Kavalan, Sameg Engi (林阿份), in June, 2012). Though some of the examples require us to modify the Base~Reduplicant correlation in (2), other examples are only apparent counterexamples to (2). These apparent counterexamples are of two types: (1) CVC-BASE that yields $\mathrm{CV_{RED}}$ rather than the expected $\mathrm{CVC_{RED}}$ (cf. (2bi)) and (2) V.CV-BASE that yields $\mathrm{V.C_{RED}}$ rather than the expected $\mathrm{V.CV_{RED}}$ (cf. (2ai)). It will be shown that the apparent counterexamples of the first type involve either monosyllabic bases or bases with the first syllable ending with a geminate and those that belong to the second type all contain identical vowels in the first two syllables of the base; therefore, they will be shown to be triggered by other effects in the language.

The rest of the paper is organized as below. Section 2 starts by providing a brief introduction to the Kavalan segmental inventory and defines the syllable structure of the language, which is crucial to an understanding of the reduplicative system in the language. Following the discussion of Kavalan syllable structure is the subsection of Kavalan reduplication where additional examples of reduplication are provided and the Base~Reduplicant correlation in (2) is revised. Section 3 provides analyses based on Coetzee's (2006) OT variation model to account for the reduplicant variations and predict the relative frequency of the variants. Section 4 examines and provides an account of apparent counterexamples. Section 5 concludes the paper.

2. Kavalan syllable structure and morphology

2.1 Kavalan syllable structure

Kavalan has 16 consonants (p, t, k, q, ʔ, β, s, z, ʁ, ɬ, ɾ, m, n, ŋ, j, w) and 4 vowels (i, u, ə, a) (Li 1982, J. Lin 1996, Chang 2000). The most canonical form of Kavalan is CVCVC. Words in Kavalan must start and end with a consonant; a glottal stop will be inserted if there is no underlying consonant present (Li 1982:481, Li & Tsuchida 2006:2). For example: /amiɬ/ → [**ʔ**amiɬ] 'bell', /sanu/ → [sanu**ʔ**] 'to speak'. In the literature, the

[4] A total of 571 items of reduplicated forms from these sources are examined. The second-hand reduplicated forms were double-checked with my Kavalan consultant whose production generally coincides with the documentation in the sources.

requirement of the presence of consonants in word-final and word-initial position could be attributed to FINAL-C (3) and INITIAL-C (4). These two constraints are adopted and are assumed to play a dominant role in Kavalan.[5] As for stress, it predictably falls on the word-final syllable in the language.

(3) FINAL-C
A word must end with a consonant.[6]

(4) INITIAL-C
A word must start with a consonant.

An important issue relevant to the present study is whether consonant clusters exist in Kavalan. Transcriptions of cc-like sequence can be seen in Kavalan literature. In Lee (2009), such a sequence is treated as tautosyllabic (cf. (1aiii)). Though cc-like sequences also appear in their *Kavalan Dictionary*, it is made clear in Li & Tsuchida (2006:4) that "there is a weak gliding vowel [ɨ̯] between each consonant cluster in the word-initial position". For example, /qman/ → [qɨ̯man] 'to eat'.[7] The occurrence of the weak vowel makes it clear that the cc-like sequence is heterosyllabic (e.g. ***qɨ̯.man*** 'to eat').[8] Li & Tsuchida restrict the appearance of *ɨ̯* to the word-initial cc-like sequence presumably because Kavalan permits a word internal coda. Therefore, the first member of the cc-like sequence in the word-internal position can be properly syllabified as the coda of the preceding syllable (e.g. *ßa**k.**siw* 'to throw away' **ßa.**kɨ̯.**siw*). The list of Kavalan canonical root forms given in Chang (2000) also supports the premise that transcriptions of cc-like sequences are heterosyllabic since none of the root forms given (including V, CV, CVC, VCVC, CVCVC, CVCCVC, CVCVCVC, CVCVCVCVC, and

[5] Notice that INITIAL-C cannot be replaced by the general constraint ONSET because a word internal onsetless syllable is permitted in Kavalan (e.g. *ß**a.u**t* 'fish').

[6] "Word" in FINAL-C as well as in INITIAL-C should be defined as a grammatical word rather than a prosodic word because the reduplicant, which will be shown in §3.1 to constitute a prosodic word, can end and/or begin without a consonant (e.g. *m-asim* 'salty' > *m-**asɨ̯**~asim* 'sour' [my field note], *qaqaʔ* 'elder siblings' > ***qa**~qaqaʔ* 'the eldest sibling' [L9:135]).

[7] The gliding vowel is not consistently marked as short in Li & Tsuchida (2006), but is always marked as short in this paper to be consistent with Li & Tsuchida's description of the phenomenon.

[8] Though Li & Tsuchida (2006:4) mention that there is a tendency to avoid homorganic consonants (e.g. **pß*, **bß*, **ßm*) in adjacent position, which seems to imply that the two consonants may be close in constituency, Li & Tsuchida also show that combinations of *mp*, *qʁ*, *ʁq* are permitted. A close examination of the chart of possible combinations of consonants given in Li & Tsuchida (2006:5) shows that many other consonant clusters that share the same place of articulation are allowed, e.g. *tɬ*, *ts*, *tr*, and *tz*.

CVCVCCVCVCV) contain complex syllable margins. As no surface cluster should occur at syllable edge in Kavalan, *COMP-M (5), which prohibits complex syllable margins, must be dominant in the language. The three constraints FINAL-C, INITIAL-C and *COMP-M are in conflict and outrank DEP-IO (6), which is against insertion, as illustrated by the tableaux in (7) and (8).

(5) *COMP-M: No complex syllable margins.

(6) DEP-IO: Output segments have input correspondents.

(7) *INITIAL-C, FINAL-C and *COMP-M must outrank DEP-IO to predict ʔ insertion*

a.

/amiɫ/ 'bell'	INITIAL-C	FINAL-C	*COMP-M	DEP-IO
☞ a. ʔa.miɫ				*
b. ʔa.miɫʔ			*!	**
c. a.miɫ	*!			

b.

/sanu/ 'to speak'	INITIAL-C	FINAL-C	*COMP-M	DEP-IO
☞ a. sa.nuʔ				*
b. ʔsa.nuʔ			*!	**
c. sa.nu		*!		

(8) *INITIAL-C, FINAL-C and *COMP-M must outrank DEP-IO to predict ɨ̯ insertion*

/qman/ 'to eat'	INITIAL-C	FINAL-C	*COMP-M	DEP-IO
☞ a. qɨ̯.man				*
b. qman			*!	

The ranking ||*COMP-M >> DEP-IO|| also correctly predicts no ɨ̯ insertion is necessary to resolve word internal cc-sequence.

(9) *No ɨ̯ insertion is necessary to resolve word internal cc-sequence*

/βaksiw/ 'to throw away'	INITIAL-C	FINAL-C	*COMP-M	DEP-IO
☞ a. βak.siw				
b. βa.kɨ̯.siw				*!

(10) summarizes the syllable structure constraints developed in this section.

(10) *Syllable structure constraints*
INITIAL-C, FINAL-C, *COMP-M >> DEP-IO

2.2 Kavalan reduplication

Listed in (11) to (14) are additional examples of Kavalan reduplication. Data cited below are accompanied by their sources. For example, data from Li & Tsuchida is cited as ‘L&T:x’, where ‘x’ is a page number, data from D. Lin as ‘DL:x’, data from J. Lin as JL:x, data from Chang as ‘C:x’, and data from Lee (2007, 2009, 2010) as ‘L7:x’, ‘L9:x’, and ‘L10:x’, respectively.

(11) *More examples of Kavalan reduplication—Consonant initial base*

a. CCV-BASE

γ1. post vocalic: CCV-BASE → C.CV$_{\text{RED}}$

i.	mu-ɬnap	‘to whisper’	>	mu-ɬ.na~ɬnap	‘to keep whispering’	[L&T:187]
ii.	βɨ̯tiq (cf. /βtiq/)	‘to go off (as of a trap)’	>	ɾa-β.ti~βtiq	‘to jump and hop’	[L&T:94]

γ2. elsewhere: CCV-BASE → Cɨ̯.CV$_{\text{RED}}$

i.	kɨ̯tun (cf. /ktun/)	‘to cut’	>	kɨ̯tu~ktun-an	‘to chop off into pieces’	[my field note]
ii.	tɨ̯mawaʁ (cf. /tmawaʁ/)	‘tomorrow’	>	pit-tɨ̯ma~tmawaʁ	‘everyday’	[L7:266]

b. CVC-BASE: → CVC$_{\text{RED}}$

i.	tumβəs	‘to pull’	>	tum~tumβəs	‘to keep pulling’	[my field note]
ii.	m-piqpiq	‘to clap’	>	piq~piqpiq	‘to keep clapping’	[my field note]
iii.	βaŋtiʔ	‘smelly’	>	su-βaŋ~βaŋtiʔ	‘to stink’	[L&T:78]

c. CV.-BASE

γ1. CV.C-BASE → CV.CV$_{\text{RED}}$

i.	βutiq-		>	ɾa-βuti~βutiq	‘to keep jumping’	[L&T:102]
ii.	su-kuɾisəŋ	‘to run around’	>	su-kuɾi~kuɾisəŋ	‘to keep running around’	[L&T:138]
iii.	pukun	‘to hit’	>	puku~pukun	‘to keep hitting’	[C:60]

γ2. CV.C-BASE → CV$_{\text{RED}}$

i.	mɨ̯-ɾizaq	‘happy’	>	mɨ̯~ɾi~ɾizaq	‘very happy’	[my field note]
ii.	ɾazat	‘person’	>	su-ɾa~ɾazat	‘to smell a lot’	[L10:105]
iii.	ma-nuβiʔ	‘to race’	>	ma-nu~nuβiʔ	‘racing’	[L&T:210]

γ3. CV.C-BASE → CVC$_{\text{RED}}$

i.	βaqiʔ	‘grandfather’	>	βaq~βaqiʔ	‘ancestors’	[L&T:78]
ii.	qatiw	‘to go’	>	qat~qatiw-an	‘the place where one often goes’	[my field note]
iii.	siʁət	‘underarm’	>	su-siʁ~siʁət	‘underarm smell’	[L10:104]

δ1. CV.V-BASE → CV.V$_{\text{RED}}$

i.	sa-βuaʔ	'to foam'	>	βua~βuaʔ	'froth at the mouth'	[L&T:95; DL:12]
ii.	mɨ̭-saiz	'cheap'	>	sai~saiz-an ta tamun ta βaɾiw	'The vegetable we sell is cheaper.'	[my field note]
iii.	suaj	'grass'	>	sua~suaj-an	'a place full of grass'	[my field note]

δ2. CV.V-BASE → CV$_{\text{RED}}$

i.	nauŋ	'mountain'	>	na~nauŋ-an	'a place of mountains'	[my field note]
ii.	βaut	'fish'	>	ʁi-βa~βaut-an	'fishing season'	[my field note]

(12) *More examples of Kavalan reduplication—Vowel initial base*

a. V.CV-BASE: → V.CV$_{\text{RED}}$

i.	m-uzan	'to rain'	>	m-uza~wzan[9]	'to keep raining'	[my field note]
ii.	m-ilam	'strong'	>	m-ila~jlam	'somewhat strong'	[my field note]
iii.	m-inaʁ	'to wake up'	>	m-ina~jnaʁ	'to wake up with lots of noise'	[my field note]

b. VC-BASE: → VC$_{\text{RED}}$

i.	ʔiβŋaw (cf. /iβŋaw/)	'spider web'	>	ʔiβ~iβŋaw-an	'a place full of spider webs'	[my field note]
ii.	ʔiʁβəŋ (cf. /iʁβəŋ/)	'deep'	>	ʔiʁ~iʁβəŋ	'very deep'	[my field note]
iii.	ʔuŋɾaj (cf. /uŋɾaj/)	'pineapple'	>	ʔuŋ~uŋɾaj-an	'a place full of pineapples'	[my field note]

(13) *Examples of CVC-BASE that unexpectedly correspond to CV$_{\text{RED}}$*

a.	ɾu-βiŋŋaw	'to become dizzy'	>	ɾu-βi~βiŋŋaw	'very dizzy'	[L&T:87]
b.	nappaw-an	'spouse'	>	na~nappaw	'to marry'	[L&T:197]
c.	qa-ɾussiq	'one'	>	qa-ɾu~ɾussiq	'one for each person'	[L&T:336]
d.	βassiŋ	'to sneeze'	>	βa~βassiŋ	'to keep sneezing'	[L&T:82]
e.	ma-βuɬ	'fist fight'	>	ma-βu~βuɬ	'fight together'	[L&T:152]
f.	sum	'urine'		su-su~sum	'smell of urine'	[L10:103]

[9] Reduplication of V.CV-BASE would result in vowel clusters across morpheme boundary; that is, V.C**V~V**.CV-BASE. Vowel clusters are not preferred in Kavalan. Gliding is often observed to repair vowel clusters (e.g. */m-uz**a~u**zan/* → *m-uz**a~w**zan* 'to keep raining' [my field note]). This repair strategy is observed not only in reduplication but also in non-reduplicated forms in the language (e.g. */m**a-i**pəs/* → *m**a-j**pəs* 'to dislike' [L&T:109]). However, exactly how vowel cluster is resolved in Kavalan involves further complications since in examples such as ***tuaβ*** 'to cough', no gliding is observed. As the study on how vowel clusters are resolved in Kavalan is premature at this stage and the investigation of the issue is out of the scope of the present paper, the vowel~glide alternation in the reduplicated forms will simply be ignored in the present study.

(14) *Examples of V.CV-BASE that unexpectedly correspond to V.C*$_{\textit{RED}}$

a.	m-ipiɾ	'to listen'	>	m-ip~ipiɾ	'to keep listening'	[my field note]
b.	m-ataʁ	'dirty'	>	m-at~ataʁ	'very dirty'	[my field note]
c.	m-isis	'to carry'	>	m-is~isis	'to keep carrying'	[my field note]
d.	m-aɾaʔ	'to take'	>	ʔaɾ~aɾa-n	'to keep taking'	[my field note]

Several things are worth noting. First, as mentioned in §2.1, a gliding vowel occurs between consonant clusters at syllable edge; furthermore, a glottal stop is inserted in word-initial and word-final position if no underlying consonant is present. These features are sometimes left unmarked in the literature due to their predictability but are marked in the examples in (11) to (14) (as well as in the examples that follow) because they play important roles in the prediction of the various reduplicant shapes. Second, since a cc-sequence at syllable edge is intervened with a gliding vowel, CCV$_{\text{RED}}$, which is considered monosyllabic in Lee (2009) (cf. (1aiii)) should be Cɨ̯.CV and disyllabic instead when in word-initial position (e.g. *kɨ̯.tu~k.tu.n-an* 'to chop off into pieces' (< *kɨ̯.tun* 'to cut')) as well as after a consonant ending prefix (e.g. *pit.-tɨ̯.ma~t.ma.waʁ* 'every day' (< *tɨ̯.ma.waʁ* 'tomorrow')). On the other hand, in word internal position after a vowel final prefix, the first consonant of CCV$_{\text{RED}}$ is syllabified as the coda of the preceding syllable (e.g. *mu-ɬ.na~ɬ.nap* 'to keep whispering' (< *mu-ɬ.nap* 'to whisper') [L&T:187]). That post-vocalic CCV$_{\text{RED}}$ is syllabified as C.CV rather than Cɨ̯.CV is predicted by the constraint ranking ||*COMP-M >> DEP-IR||, which predicts that insertion of the weak gliding vowel in the reduplicant is only necessary to repair the *COMP-M violation, as illustrated in (15).[10] Hereafter, CCV$_{\text{RED}}$ is referred to as C.CV$_{\text{RED}}$ post-vocalically and as Cɨ̯.CV$_{\text{RED}}$ elsewhere.

(15) *||*COMP-M >> DEP-IR|| predicts the syllabification of C.CV*$_{\textit{RED}}$ *after vowel*

/mu-RED~ɬnap/ 'to keep whispering'	*COMP-M	DEP-IR
☞ a. mu-ɬ.na~ɬ.nap		
b. mu.-ɬɨ̯.na~ɬ.nap		*!
c. mu.-ɬna~ɬ.nap	*!	

Third, though examples in (11) and (12) generally conform to Lee's (2009) observation given in (2), counterexamples can also be found. In particular, CV.V-BASE is found to yield various reduplicant shapes, as does CV.C-BASE. As shown in (11cδ1-2),

[10] Notice that DEP-IR functions to rule out Cɨ̯.CV$_{\text{RED}}$ in post-vocalic position, but not in other positions because a copy of the CCV-BASE without vowel insertion in non-post-vocalic position would result in complex onset (e.g. *[(*mɾi*$_{\mu}$)]~*mɾimək* < /*RED~mɾimək*/ 'taciturn' [L&T:325]), violating undominated *COMP-M, which outranks DEP-IR.

CV.V-BASE can not only yield CV.V$_{\text{RED}}$, but also CV$_{\text{RED}}$, though the latter is less frequently observed. The correlation between the structure of base initial syllable and reduplicant shape is revised in (16). (16) also provides the frequency of the different reduplicant shapes for those bases that yield variants. The reduplicant shape that is more frequently observed for each base type is marked in bold.

(16) *Correlation between base type and reduplicant shape* (*revision*)

a. Consonant initial base

γ1. CCV-BASE:

- i. post-vocalic → **C.CV$_{\text{RED}}$**
- ii. elsewhere → **Cɨ̯.CV$_{\text{RED}}$**

γ2. CVC-BASE: → **CVC$_{\text{RED}}$**

γ3. CV.-BASE:

i.	CV.C-BASE →	**CV$_{\text{RED}}$**	**(116/236 = 49%)**
		CVC$_{\text{RED}}$	(93/236 = 39%)
		CV.CV$_{\text{RED}}$	(27/236 = 11%)
ii.	CV.V-BASE →	**CV.V$_{\text{RED}}$**	**(30/42 = 71%)**
		CV$_{\text{RED}}$	(12/42 = 29%)

b. Vowel initial base

γ1. V.CV-BASE: → **V.CV$_{\text{RED}}$**

γ2. VC-BASE: → **VC$_{\text{RED}}$**

(13) and (14) contain additional examples that do not conform to the Base~Reduplicant in (2) or (16). These examples involve CVC-BASE that yields CV$_{\text{RED}}$ rather than the expected CVC$_{\text{RED}}$ and V.CV-BASE that yields V.C$_{\text{RED}}$ rather than the expected V.CV$_{\text{RED}}$. It will be shown in §4 that these examples are triggered by other effects in the language and are merely apparent rather than real counterexamples.

Despite the fact that the reduplicants of Kavalan reduplication vary in shape, they always copy from the left edge of the base. This suggests that ANCHOR-BR-L (17) is dominant in the language, as illustrated in (18). For simplicity, in the OT analysis that follows, this dominant constraint will be omitted and output candidates that violate the constraint will not be considered.

(17) ANCHOR-BR-L
The left peripheral segment of a reduplicant corresponds to the left peripheral segment of the base.

(18) ANCHOR-BR-L
ɾa~ɾazat ‘persons’ (< ɾazat ‘person’) [L&T:320]
/RED~ɾazat/
ɾa~ɾazat > za~ɾazat

3. An OT analysis

This section provides an OT analysis of Kavalan reduplication.

3.1 The reduplicant size

The reduplicant in Kavalan takes several distinct shapes and varies from monosyllabic to disyllabic and from open to closed syllables. Reduplicants with various shapes, though less common, are not unique to Kavalan. For languages that have various reduplicant shapes, there is usually a default shape, with the other variants driven by other effects. For instance, the durative reduplication in Tawala (Ezard 1997, Hicks-Kennard 2004), an Austronesian language spoken in the Milne Bay area of Papua New Guinea, has three reduplicant variants, a foot size CV.CV (e.g. *ge.le.ta* 'arrive' > *ge.le~ge.le.ta* 'be arriving'), a monosyllabic CV (e.g. *be.i.ha* 'search' > *bi.~be.i.ha* 'be searching'), and a reduplicant V.C, which is bigger than a syllable but smaller than a foot (e.g. *a.pu* 'bake' > *a.p~a.pu* 'be baking'). According to Hicks-Kennard (2004), the default size of the reduplicant is disyllabic and the other variants of the reduplicant (i.e. CV and V.C) are driven by an OCP effect, which prevents the reduplicant and the stem from having identical adjacent syllables.

With regard to the default shape of Kavalan reduplication, if we focus on the dominant reduplicant shapes of each base type (marked in bold in (16)), we can find that the reduplicant is more likely to be composed of two open syllables (i.e. Cɨ̯.CV$_{\text{RED}}$, CV.V$_{\text{RED}}$, and V.CV$_{\text{RED}}$) or a string composed of at least a vowel and a coda (i.e. CVC$_{\text{RED}}$, C.CV$_{\text{RED}}$, and VC$_{\text{RED}}$). Disyllabic reduplicants with open syllables are undoubtedly bimoraic. Reduplicants with a vowel and a coda could be too, if codas are moraic in Kavalan.

Codas in Kavalan are indeed moraic. Recall that Kavalan stress predictably falls on the last syllable and that words must end with a coda. The reason a glottal stop is inserted in the coda position if no underlying consonant is present could be to turn a light syllable heavy to make it capable of bearing stress. (19) further supports that the Kavalan coda is moraic. The examples in (19) are from Li (1982) who remarks that "the stem- and word-final glottal stop disappears when immediately followed by a suffix and word respectively." Examples in (19) portray a clear correlation between the presence of coda and stress location: a stressed syllable always carries a coda and an inserted glottal stop is lost when a syllable is no longer stressed. Therefore, Kavalan codas are moraic.

(19) *Correlation between coda and stress location* (*Li 1982:481*)
a. [mʁijaː**sáʔ**] ‘to buy’,
b. [ʁijas**a**-i**káʔ**] ‘Buy!’,
c. [mʁijas**a**-iku ɬaː**kán**] ‘I bought something’,
d. [βaqijan mʁijas**a** tu βaː**út**] ‘The old man bought some fish’.

The correlation between the presence of coda and stress also suggests that word-final glottal stop insertion in Kavalan, which was previously attributed to FINAL-C, should actually be triggered by STRESS-TO-WEIGHT, which requires all stressed syllables to be heavy (Prince 1990, Yu 2005). STRESS-TO-WEIGHT is undominated in Kavalan since there is no surface exception. As STRESS-TO-WEIGHT is satisfied by glottal stop insertion rather than vowel lengthening, DEP-IO must be outranked by *LONG-V, as illustrated in (22). The syllable structure constraints of (10) are revised in (23).

(20) STRESS-TO-WEIGHT (SWP) (Yu 2005)
Stressed syllable must be heavy.

(21) *LONG-V: Long vowels are prohibited.

(22) *SWP and *LONG-V must outrank DEP-IO*

/kama/ ‘orange (loan)’	SWP	*LONG-V	DEP-IO
☞ a. kamáʔ			*
b. kamáː		*!	
c. kamá	*!		

(23) *Syllable structure constraints* (*revision*)
INITIAL-C, *COMP-M, SWP, *LONG-V >> DEP-IO

The moraic status of the coda suggests that, just as with the disyllabic reduplicant of open syllables, reduplicants with a vowel and a coda (i.e. CVC_{RED}, $C.CV_{RED}$, and VC_{RED}) are bimoraic as well (i.e. $CV_{\mu}C_{\mu}$, $C_{\mu}.CV_{\mu}$, and $V_{\mu}C_{\mu}$). Thus, the default reduplicant size of Kavalan reduplication is assumed to be bimoraic.

Following McCarthy & Prince (1990, 1994a, 1994b) and their followers such as Kager (1999), Hendricks (1999), and Crowhurst (2004), the bimoraic shape of the reduplicant is analyzed as a prosodic word in the present study. The two constraints in (24) and (25) are proposed. By assuming that a prosodic word must dominate a foot (i.e. with undominated HEADNESS constraint) and that a foot is bimoraic (i.e. FTBIN), a reduplicant will be minimally bimoraic.[11]

[11] McCarthy & Prince (1990, 1994a, 1994b) propose that any given reduplicant is specified underlyingly either as an affix or a stem. Given the two general constraints AFFIX ≤ σ and STEM = PRWD that impose size restriction on affixes and stems that are independent of reduplicant, a reduplicant will be no larger than a syllable and no smaller than a minimal word. In Kavalan,

(24) RED-PRWD-L
Align the left edge of a RED with the left edge of a prosodic word.

(25) RED-PRWD-R
Align the right edge of a RED with the right edge of a prosodic word.

RED-PRWD-L and RED-PRWD-R together only require a reduplicant to be minimally, but not maximally, bimoraic since RED-PRWD-L and RED-PRWD-R together only require the left and the right edges of the reduplicant to be aligned with the left and right edges of some prosodic word; thus, candidates such as [(<u>μμ</u>)]~BASE (where […] = prosodic word, (…) = foot), satisfies both constraints as well as [(<u>μμ</u>)][(<u>μμ</u>)]~BASE, even though they contain a different number of prosodic words. DEP-OO-SEG (Benua 1997), which is an OO-faithfulness constraint that requires every segment in the reduplicated form to have a correspondent in its corresponding unreduplicated form, can help limit the reduplicant to its minimal size (Gouskova 2003, 2004).[12] By ranking it below RED-PRWD-L and RED-PRWD-R and above constraints which favor copying (e.g. MAX-BR), the reduplicant is limited to being bimoraic, as illustrated in (27).

(26) DEP-OO-SEG (Benua 1997, Gouskova 2003, 2004)
Every segment in the reduplicated form has a correspondent in its corresponding unreduplicated form [henceforth, R(eference) O(utput)].

(27) *DEP-OO-SEG must outrank MAX-BR*

/RED~tuktuk/ ‘to keep knocking’ [L9:135] RO: tuktuk ‘to knock’	RED-PRWD-L	RED-PRWD-R	DEP-OO-SEG	MAX-BR
☞ a. [(<u>$tu_{\mu}k_{\mu}$</u>)]~tuktuk			t, u, k	***
b. [(<u>$tu_{\mu}k_{\mu}$</u>)][(<u>$tu_{\mu}k_{\mu}$</u>)]~tuktuk			t, u, k, t,! u, k	

there is no metrical evidence to support the idea that the reduplicant forms a prosodic word at this moment because the study on the metrical structure of Kavalan is still premature and it is only clear at this moment that stress falls in word-final position. But sometimes a reduplicant is analyzed as a prosodic word without metrical evidence in the literature. For instance, the reduplicant in Makassarese and in Kamaiura is analyzed as a prosodic word in McCarthy & Prince (1994b) and in Crowhurst (2004), respectively, without support from metrical structure in the respective languages. They are analyzed as prosodic words because the reduplicant in both languages is bimoraic in size.

12 A common a-templatic analysis that limits the size of the reduplicant is to rank some structure penalizing constraints (e.g. *STRUC) above constraints which favor copying (Gafos 1998, Kennedy 2005, Spaelti 1997, and Yu 2005). However, Gouskova (2003, 2004) criticizes that such an approach has the disadvantage of predicting languages in which even stems are limited to a minimum size, which are unattested. I am grateful to an anonymous reviewer for pointing this out.

As a matter of fact, DEP-OO-SEG will not only limit the reduplicant to being bimoraic but also prefers a tri-segmental reduplicant with coda to a quadri-segmental reduplicant with two open syllables because each additional segment copied will incur an additional violation in DEP-OO-SEG, as illustrated in (28).

(28) *DEP-OO-SEG prefers monosyllabic CVC$_{RED}$ to disyllabic CV.CV$_{RED}$*[13]

RO: BASE	DEP-OO-SEG
☞ a. [($\underline{CV_{\mu}C_{\mu}}$)]~BASE	***
b. [($\underline{CV_{\mu}CV_{\mu}}$)]~BASE	****!

Thus, the constraint easily predicts CVC$_{RED}$ for CVC-BASE (cf. (28)). DEP-OO-SEG also easily predicts a post-vocalic reduplicant corresponding to CCV-BASE to be the shorter C.CV$_{RED}$ rather than the longer Cɨ̯.CV$_{RED}$, as illustrated in (29).

(29) *C.CV$_{RED}$ is correctly predicted for CCV-BASE in post-vocalic position*

/sa-RED~mʁiŋu/ ‘to pretend not to know’ [L&T:351] RO: mɨ̯ʁiŋuʔ ‘to have no idea’	DEP-OO-SEG	MAX-BR
☞ a. sa-[(m$_\mu$.ʁi$_\mu$)]~m.ʁi.ŋuʔ	s, a, m, ʁ, i	***
b. sa.-[(mɨ̯$_\mu$.ʁi$_\mu$)]~m.ʁi.ŋuʔ	s, a, m, ɨ̯, ʁ,! i	***

Notice that though both CVC$_{RED}$ and C.CV$_{RED}$ are favored by the current constraint ranking, C.CV$_{RED}$ can only correspond to CCV-BASE. This is predicted by the LINEARITY-IR constraint (30), which is assumed to play a dominant role in the language because no surface exception is found. Consider (31) below.

(30) LINEARITY-IR
No segment reversal in input-reduplicant.

(31) *LINEARITY-IR predicts C.CV$_{RED}$ can only be yielded by CCV-BASE*

/ma-RED~$C_1V_2C_3$CVC/	LINEARITY-IR
☞ a. ma-[($\underline{C_1V_2C_3}$)]~$C_1V_2C_3$CVC	
b. ma-[($\underline{C_1.C_3V_2}$)]~$C_1V_2C_3$CVC	*!

[13] Notice that the candidate that best satisfies the DEP-OO-SEG constraint is a candidate without a reduplicant. Such a candidate violates REALIZE MORPHEME (Kurisu 2001 and others) which requires a morpheme to have some phonological exponent in the output. REALIZE MORPHEME is undominated in Kavalan. For simplicity, this dominant constraint will be omitted in the OT analyses that follow and output candidates that violate the constraint will not be considered.

However, for non-post-vocalic reduplicants that correspond to CCV-BASE, DEP-OO-SEG would wrongly predict the reduplicant to be Cɨ̯C$_{RED}$ rather than the attested Cɨ̯.CV$_{RED}$ because the former contains fewer segments and thus incurs fewer violations of DEP-OO-SEG. Cɨ̯C$_{RED}$ could be easily ruled out, however, because the reduplicant, which takes the form of a foot, is headed by the high central gliding vowel *ɨ̯*. There is a universal preference for a foot to be headed by vowels of higher sonority (Kenstowicz 1997, de Lacy 2004). According to the vowel-sonority hierarchy in (32), a high central vowel has the least sonority. Thus, it is marked for *ɨ̯*, which is not only a high central vowel, but is also gliding, to head a foot. The *FT/ɨ̯ constraint in (33) is thus proposed to prohibit the situation. When it is ranked above DEP-OO-SEG, the reduplicant is correctly predicted as Cɨ̯.CV$_{RED}$, as illustrated in (34).

(32) *Vowel-sonority hierarchy* (*de Lacy 2004*)
low peripheral > mid peripheral > high peripheral > mid central > high central
‘a’ ‘e.o’ ‘i.u’ ‘ə’ ‘ɨ’

(33) *FT/ɨ̯
The only vowel of the foot cannot be the weak gliding *ɨ̯*.

(34) *||*FT/ɨ̯ >> DEP-OO-SEG|| prefers disyllabic Cɨ̯.CV$_{RED}$ to monosyllabic Cɨ̯.C$_{RED}$*

RO: BASE	*FT/ɨ̯	DEP-OO-SEG
☞ a. [(Cɨ̯.CV)]~BASE		****
b. [(Cɨ̯C)]~BASE	*!	***

(35) summarizes the constraint ranking developed in this section to predict the default reduplicant shape.

(35) *Constraint hierarchy to predict the reduplicant shape*
HEADNESS, RED-PRWD-L, RED-PRWD-R, *FT/ɨ̯, FTBIN, LINEARITY-IR
>> DEP-OO-SEG
>> MAX-BR

3.2 Reduplication in consonant initial bases

This section discusses reduplication in consonant initial bases, which includes CV.C-BASE, CV.V-BASE, CVC-BASE, and CCV-BASE.

3.2.1 The various reduplicant shapes

In §3.1, it is proposed that the default reduplicant is bimoraic in size and prefers to be as small as possible; thus a tri-segmental reduplicant with a vowel and a coda is generally favored over a quadri-segmental reduplicant with two open syllables. However, some base types either do not yield the preferred reduplicant shape or yield other reduplicant shapes as well. These base types are considered below.

3.2.1.1 Reduplicant variants of CV.C-BASE

Since the best reduplicant shape is the smallest bimoraic unit, for CV.C-BASE, the reduplicant is naturally predicted to be CVC$_{\text{RED}}$. However, CVC$_{\text{RED}}$ is only one of the variants yielded by CV.C-BASE; CV.C-BASE also yields CV$_{\text{RED}}$ and CV.CV$_{\text{RED}}$.[14] As CVC-BASE yields only CVC$_{\text{RED}}$, why CVC$_{\text{RED}}$ is not good enough to stand as the only reduplicant for CV.C-BASE remains puzzling.

McCarthy (2000) proposes that in addition to segmental faithfulness, there should also be prosodic faithfulness constraints requiring that certain properties of prosodic structures such as feet and syllables be preserved in related forms. Prosodic faithfulness, as proposed in McCarthy, is mediated by the edge or the head of the constituent and is captured by the Anchoring constraint.[15] Prosodic faithfulness provides an answer to the various reduplicant shapes of CV.C-BASE. Clearly, while the syllable structure of CVC$_{\text{RED}}$ yielded by CVC-BASE faithfully matches the underlying syllable structure of the base (cf. (36a)), no such prosodic faithfulness exists between the syllable structure of CVC$_{\text{RED}}$ yielded by CV.C-BASE and the underlying structure of the base since the coda of CVC$_{\text{RED}}$ corresponds to an underlying base onset, but not to a coda (cf. (36b)).[16]

(36) *IR correspondence of syllable structure*

a. C V C$_{\text{RED}}$ /C V C/-BASE vs. b. C V C$_{\text{RED}}$ /C V. C/-BASE

[14] The choice among different variants of CV.C-BASE is not phonologically governed. Though it is mentioned in Lee (2007, 2009) that there is a tendency for CVC$_{\text{RED}}$ to occur when the coda corresponds to a nasal (or glide) onset in the base (e.g. *sunis* 'child' > sun~*sunis* 'children' [L9:137]) and that CV$_{\text{RED}}$ is preferred to CVC$_{\text{RED}}$ if copying an additional segment would result in identical adjacent sequence (e.g. *sisəp* 'to suck' > si~*sisəp* 'to keep sucking' *sis~*sisəp* [my field note]), many examples contrary to the tendency can be found. For example, there are many CVC$_{\text{RED}}$ that do not end with a nasal or glide (e.g. *maʁaw* 'to look' > maʁ~*maʁaw* 'to keep looking' [L&T:158]). This is probably why Lee (2009:136) also remarks that "there seems to be no specific rules as to which reduplicant will be yielded."

[15] Evidence in support of syllable faithfulness is reported in Yu (2007), Piñeros (1998), and H. Lin (2010) while evidence of foot faithfulness is available in McCarthy (1995, 2000).

[16] McCarthy (2000) remarks that under the principle of the *richness of the base*, prosodic structures (e.g. syllable) are allowed to be present in the underlying representation (cf. also Alderete 1996, McCarthy 1995, Inkelas 1999, Itô et al. 1996).

In other words, a CVC$_{RED}$ yielded by CV.C-BASE would violate the prosodic faithfulness constraint ANCHOR-IR$_σ$ defined in (37) (cf. Piñeros 1998, Yu 2007, H. Lin 2010), while the same reduplicant yielded by CVC-BASE would not. For CV.C-BASE, the perfect CVC shape would incur a violation of ANCHOR-IR$_σ$, but an undersized CV$_{RED}$ would have a perfect syllable match between the base and the reduplicant and would satisfy ANCHOR-IR$_σ$. Compare (38) with (39) below. Thus, ANCHOR-IR$_σ$ provides an explanation as to why the CV$_{RED}$ variant is yielded by CV.C-BASE but not by CVC-BASE.

(37) ANCHOR-IR$_σ$
The initial and final position of two syllables in an Input-Reduplicant correspondence relationship must correspond.

(38) *CVC$_{RED}$ yielded by CVC-BASE fully satisfies ANCHOR-IR$_σ$*

/RED~tan.-βa.səʁ/ 'to keep flying off' [my field note]	ANCHOR-IR$_σ$
☞ a. tan.~tan.-βa.səʁ	
b. ta.~tan.-βa.səʁ	*!

(39) *CV$_{RED}$ yielded by CV.C-BASE fully satisfies ANCHOR-IR$_σ$*

/RED~ɾa.tə.βus/ 'very sweet' [my field note]	ANCHOR-IR$_σ$
☞ a. ɾa.~ɾa.tə.βus	
b. ɾat.~ɾa.tə.βus	*!

Though ANCHOR-IR$_σ$ helps predict CV$_{RED}$ for CV.C-BASE, CV$_{RED}$ is again not the only variant yielded by CV.C-BASE; CV.C-BASE can also yield CVC$_{RED}$ and CV.CV$_{RED}$. In a constraint-reranking approach, the alternation between CVC$_{RED}$ and the undersized CV$_{RED}$ can be attributed to the ranking differences between FTBIN and ANCHOR-IR$_σ$; ||FTBIN >> ANCHOR-IR$_σ$|| predicts the CVC$_{RED}$ variant, while the reverse ranking predicts the CV$_{RED}$ variant, as illustrated in (40) and (41), respectively.

(40) *||FTBIN >> ANCHOR-IR$_σ$|| predicts the CVC$_{RED}$ variant for CV.C-BASE*

/RED~si.nap/ 'to sweep everywhere' [my field note]	FTBIN	ANCHOR-IR$_σ$
☞ a. [(si$_μ$n$_μ$.)]~si.nap		*
b. [(si$_μ$.)]~si.nap	*!	

(41) *||ANCHOR-IR$_σ$ >> FTBIN|| predicts the CV$_{RED}$ variant for CV.C-BASE*

/RED~ɾa.tə.βus/ 'very sweet' [my field note]	ANCHOR-IR$_σ$	FTBIN
☞ a. [(ɾa$_μ$.)]~ɾa.tə.βus		*
b. [(ɾa$_μ$t$_μ$.)]~ɾa.tə.βus	*!	

Additionally, to predict the variation between CV.CV$_{RED}$ and CVC$_{RED}$, one can propose a re-ranking between DEP-OO-SEG, which least prefers CV.CV$_{RED}$, and constraints such as NOCODA, which disfavors CVC$_{RED}$. However, the constraint re-ranking approach fails to capture the fact that though CV.CV$_{RED}$ is a possible variant for a CV.C-BASE, it is not as common as the other two variants. For the relative frequency of different variants, please refer to (16).

Coetzee (2006) proposes a Rank-Ordering Model of EVAL (ROE) which not only accounts for variations but also predicts the relative frequency of the variants. Unlike approaches that depend on constraint re-ranking, variation in ROE does not result from variation in grammar (i.e. ranking). Instead, variation depends on how EVAL imposes a well-formedness ranking order on the candidates. In the model, which allows only one constraint ranking, a 'critical cut-off' separates the constraints into two strata: those that ranked higher than the cut-off and those that ranked lower than the cut-off. Only violations of constraints above the cut-off are fatal, just as in classic OT (where non-optimal candidates do not surface); violations of constraints below the cut-off are not severe enough to rule out candidates. All candidates that survive upon reaching the cut-off are grammatical output variants. The relative degrees of well-formedness of the candidates that pass through the cut-off indicate the relative frequency of the variants. The illustrative tableau from Coetzee is given in (42).

(42) *The Ranking-Ordering Model of EVAL* (*Coetzee 2006:343*)

cut-off

	ℂ1	ℂ2	ℂ3	ℂ4
☞$_1$ a. Cand$_1$				*
☞$_2$ b. Cand$_2$			*	
c. Cand$_3$		*!		
d. Cand$_4$	*!			

In (42), candidates (c) and (d) have fatal violations above the cut-off and are ruled out. Violations of the constraints below the cut-off (which is indicated by a thick vertical line) are not fatal. Thus, both candidates (a) and (b) are grammatical. Both candidates (a) and (b) incur one violation below the cut-off; however, candidate (a) is more well-formed than (b) because it violates a lower ranked constraint. Thus, it is predicted to be observed more frequently than candidate (b). (In the tableau, the pointing hand indexed with subscript 1 points to the most commonly observed variant and the pointing hand indexed with subscript 2 points to the second most commonly observed variant, etc.) In ROE, the constraints violated by the output variants must be ranked below the cut-off to ensure the variants do surface.

In Kavalan, the variants of CV$_{\text{RED}}$, CVC$_{\text{RED}}$ and CV.CV$_{\text{RED}}$ yielded by CV.C-BASE violate at least one of the constraints in (43).

(43) *Constraints violated by CV$_{\text{RED}}$, CVC$_{\text{RED}}$ and CV.CV$_{\text{RED}}$ corresponding to CV.C-BASE*

	FTBIN	ANCHOR-IR$_\sigma$	DEP-OO-SEG
CVC$_{\text{RED}}$		*	***
CV.CV$_{\text{RED}}$		(*)	****
CV$_{\text{RED}}$	*		**

CVC$_{\text{RED}}$ has the perfect bimoraic shape and thus satisfies FTBIN. It has three segments and violates DEP-OO-SEG three times. It also violates ANCHOR-IR$_\sigma$ because of the mismatch of the syllable structures between the reduplicant and the underlying structure of the base. CV.CV$_{\text{RED}}$ also has the perfect bimoraic shape and satisfies FTBIN. It may or may not obey ANCHOR-IR$_\sigma$ (e.g. /RED~*ka.si.-a.nəm*/ → [<u>*ka.si.*</u>~*ka.si.-a.nəm*] 'to be thinking about' [L7:176], /*pi.-RED~ɾi.paj*/ → [*pi.-*<u>*ɾi.pa.*</u>~*ɾi.pa.j-an*] 'every Sunday' [my field note]), but it incurs more violations of the DEP-OO-SEG constraint than does CVC$_{\text{RED}}$. CV$_{\text{RED}}$ is monomoraic. Due to top-ranked RED-PRWD-L, RED-PRWD-R, and HEADNESS, it is considered to form a prosodic word and, therefore, a foot as well. However, since the foot is degenerate, it violates FTBIN. But as there is a perfect matching of the syllable structure of the reduplicant to the base input, it fully satisfies ANCHOR-IR$_\sigma$. It also best satisfies DEP-OO-SEG because it contains the fewest segments.

Because CV$_{\text{RED}}$, CVC$_{\text{RED}}$ and CV.CV$_{\text{RED}}$ violate at least one of the constraints in (43), these constraints must be ranked below the cut-off. Furthermore, since CV.CV$_{\text{RED}}$ is less common, the constraint that least favors it must be ranked higher below the cut-off. (44) illustrates that the ranking of ||DEP-OO-SEG >> FTBIN >> ANCHOR-IR$_\sigma$|| below the cut-off successfully predicts the reduplicant variants (i.e. CV$_{\text{RED}}$, CVC$_{\text{RED}}$ and CV.CV$_{\text{RED}}$) for CV.C-BASE and the relevant frequency of the variants. However, because the three constraints are ranked below the cut-off and are thus non-fatal, it will wrongly predict CV.CVC$_{\text{RED}}$ as a possible reduplicant variant for CV.C-BASE, which just does not occur as frequently as the rest of the candidates (ref. (44d)).

(44) *Frequency of occurrence predicted by ROE*

CV.C-BASE	C1	DEP-OO-SEG	FTBIN	ANCHOR-IR$_\sigma$
☞$_1$ a. CV$_{\text{RED}}$		**	*	
☞$_2$ b. CVC$_{\text{RED}}$		***		*
☞$_3$ c. CV.CV$_{\text{RED}}$		****		(*)
💣$_4$ d. CV.CVC$_{\text{RED}}$		*****	*	(*)

CV.CVC$_{\text{RED}}$ is trimoraic and oversized (i.e. CV$_\mu$.CV$_\mu$C$_\mu$). But FTBIN is ranked below the cut-off to predict the undersized variant CV$_{\text{RED}}$. In Kavalan reduplication, while the reduplicant can be smaller than bimoraic, it is never bigger than two moras. Thus, following Everett (1996, 2003), Downing (2000), Chen (2000), and Hsiao (2000, 2008), among others, this paper reformulates FTBIN as FTMAX and FTMIN below:

(45) FTMAX
Feet are maximally bimoraic.

(46) FTMIN
Feet are minimally bimoraic.

Since the reduplicant is never over-sized, FTMAX must be undominated. FTMIN, on the other hand, has to be ranked below the cut-off to predict the surface of CV_{RED} for CV.C-BASE, as illustrated in (47). In (47), candidate (d) is ruled out as a possible candidate by FTMAX which is ranked above the cut-off. Candidate (a)-(c) survive upon reaching the cut-off and are grammatical outputs. Judging from the violation of the top-ranked constraint below the cut-off (i.e. DEP-OO-SEG) (a) is the most harmonic candidate and (c) the least harmonic; therefore, (a) is correctly predicted as the most commonly observed variant and (c) the least frequently observed one.

(47) *FTMAX and FTMIN help distinguish undersized and oversized reduplicants*

CV.C-BASE	FTMAX	DEP-OO-SEG	FTMIN	ANCHOR-IR$_\sigma$
☞1 a. CV_{RED}		**	*	
☞2 b. CVC_{RED}		***		*
☞3 c. $CV.CV_{RED}$		****		(*)
d. $CV.CVC_{RED}$	*!	*****		(*)

Before ending this section, it is worth noting that the distinct behavior of CVC_{RED} corresponding to CVC-BASE and CV.C-BASE could also have been attributed to the BR matching of syllable structure (i.e. ANCHOR-BR$_\sigma$) since the syllable structure of CVC_{RED} corresponding to CVC-BASE also faithfully matches the syllable structure of the base output while the syllable structure of CVC_{RED} yielded by CV.C-BASE also fails to match the syllable structure of its corresponding base output. Nonetheless, the examples in (48) show that it is the underlying (rather than the surface) syllable structure of the base that matters.

(48) *Bases with initial syllable as CVC but reduplicant as CV_{RED}*

a.	qajnəp	'to sleep'	>	qa~qajnəp	'to keep sleeping'	[JL:75]
b.	siwpan	'to puff air'	>	si~siwpan	'to keep puffing air'	[JL:75]
c.	ziwnan	'to shake'	>	zi~ziwnan	'to keep shaking'	[JL:75~76]

In (48), CVC-BASE unexpectedly yields CV_{RED}. This could never be explained by ANCHOR-BR_{σ} since the syllabic matching between (the unattested) CVC_{RED} and the base output (i.e. CVC) is better than between (the attested) CV_{RED} and the base output.

(49) *ANCHOR-BR_{σ} fails to predict examples in* (*48*)
e.g. *qajnəp* 'to sleep' > *qa~qajnəp* 'to keep sleeping'
qaj.~**qaj**.nəp > **qa**.~**qaj**.nəp
BR correspondence

The puzzle can be solved by looking into the underlying syllable structure of the bases. The roots of the examples, as listed in the *Kavalan Dictionary* are *inəp*-, *siup*-, and *ziun*-, respectively. Thus, the glide codas in the first syllable of the base output are actually vowels underlyingly (which turn into glides presumably to resolve vowel clusters). Therefore, the underlying syllable structure of the base is CV.V rather than CVC. Thus, ANCHOR-IR_{σ} (but not ANCHOR-BR_{σ}) correctly predicts the surface of CV_{RED} since the syllabic matching between CV_{RED} and the underlying structure of the base is better than CVC_{RED} and the underlying structure of its base.[17]

(50) *ANCHOR-IR_{σ} makes the correct prediction for* (*48*)
/**qa.**-i.nəp/ /**qa**.-i.nəp/
| IR correspondence |
qa.~qaj.nəp > **qaj.**~qaj.nəp

3.2.1.2 Reduplicant variants of CV.V-BASE

When dealing with CV.C-BASE, it is proposed that the three constraints, ANCHOR-IR_{σ}, DEP-OO-SEG and FTBIN, must be ranked below the cut-off to predict the various reduplicant shapes. Just like CV.C-BASE, CV.V-BASE, which starts with a syllable of CV also yields reduplicant variants: $CV.V_{RED}$ and CV_{RED}, though the latter is less frequently observed. The constraints that are placed below the cut-off to yield the variants for CV.C-BASE can also predict the two reduplicant variants for CV.V-BASE, as illustrated in (51), except that the ranking ||DEP-OO-SEG >> FTMIN >> ANCHOR-IR_{σ}||, which predicts the higher frequency of monosyllabic CV_{RED} over disyllabic $CV.CV_{RED}$ for CV.C-BASE would also predict that CV.V-BASE would yield monosyllabic CV_{RED} more often than disyllabic $CV.V_{RED}$, which is counterfactual. Consider (51).

[17] Notice that the need for IR correspondence is not limited to Kavalan reduplication. Please refer to McCarthy & Prince (1995) for why IR correspondence is necessary as well as the typological consequences of IR correspondence.

(51) *||DEP-OO-SEG >> FTMIN >> ANCHOR-IR$_{\sigma}$|| makes the wrong prediction of frequency*

CV.V-BASE	DEP-OO-SEG	FTMIN	ANCHOR-IR$_{\sigma}$
💣$_1$ a. CV$_{\text{RED}}$	**	*	
💣$_2$ b. CV.V$_{\text{RED}}$	***		(*)

We are facing a dilemma: while CV.C-BASE prefers a monosyllabic reduplicant to a disyllabic one, the preference of CV.V-BASE is the opposite. This results in a ranking paradox between DEP-OO-SEG and FTMIN. (52i) shows that FTMIN must outrank DEP-OO-SEG to correctly predict the frequency of the reduplicant variants for CV.V-BASE; however, as illustrated in (52ii), to make the correct prediction for CV.C-BASE, FTMIN must be dominated (or equally ranked) with DEP-OO-SEG.

(52) *Ranking paradox between DEP-OO-SEG and FTMIN*

i.

CV.V-BASE	FTMIN	DEP-OO-SEG
☞$_1$ a. CV.V$_{\text{RED}}$		***
☞$_2$ b. CV$_{\text{RED}}$	*	**

ii.

CV.C-BASE	DEP-OO-SEG	FTMIN
☞$_1$ a. CV$_{\text{RED}}$	**	*
☞$_2$ b. CV.CV$_{\text{RED}}$	****	

The dilemma can be solved if DEP-OO-SEG is divided into DEP-OO-V and DEP-OO-C. In the CV.V$_{\text{RED}}$~CV$_{\text{RED}}$ variation, CV.V$_{\text{RED}}$ and CV$_{\text{RED}}$ have the same number of consonants, but CV.V$_{\text{RED}}$ has one more vowel than CV$_{\text{RED}}$ does. In the CV.CV$_{\text{RED}}$~CV$_{\text{RED}}$ variation, CV.CV$_{\text{RED}}$ not only outnumbers CV$_{\text{RED}}$ in vowels, but also in consonants. In other words, what truly disprefers CV.CV$_{\text{RED}}$ to CV$_{\text{RED}}$ is the additional consonant, not the vowel, in CV.CV$_{\text{RED}}$. Thus, when DEP-OO-C outranks FTMIN, which in turn outranks DEP-OO-V, the correct frequency can be predicted for both CV.C-BASE and CV.V-BASE, as illustrated in (53). Take (53ii) for instance; both candidates (a) and (b) violate DEP-OO-C once because they both contain a consonant in the reduplicant; however, since (b) is monomoraic, it incurs an additional violation in FTMIN. Consequently, candidate (a), which is better formed than (b), is correctly predicted as being more frequently observed.

(53) *Distinction of DEP-OO-C and DEP-OO-V solves the ranking paradox*

i.

CV.C-BASE	DEP-OO-C	FTMIN	DEP-OO-V	ANCHOR-IR$_{\sigma}$
☞$_1$ a. CV$_{\text{RED}}$	*	*	*	
☞$_2$ b. CV.CV$_{\text{RED}}$	**		**	(*)

ii.

CV.V-BASE	DEP-OO-C	FTMIN	DEP-OO-V	ANCHOR-IR$_\sigma$
☞$_1$ a. CV.V$_{RED}$	*		**	(*)
☞$_2$ b. CV$_{RED}$	*	*	*	

3.2.2 Lack of variation in CVC-BASE and CCV-BASE

To predict the reduplicant variants for CV.C-BASE and CV.V-BASE, ANCHOR-IR$_\sigma$, DEP-OO-V, DEP-OO-C, and FTMIN are ranked below the cut-off. It is important to find out whether they will make the wrong prediction for CVC-BASE and CCV-BASE, which do not yield reduplicant variants. In particular, since FTMIN is placed below the cut-off, a monosyllabic reduplicant is likely to be wrongly predicted as a variant for CVC-BASE and CCV-BASE as well.

3.2.2.1 CVC-BASE

Indeed, for CVC-BASE, the constraint ranked below the cut-off not only predicts CV$_{RED}$ as a possible variant but also predicts it would be observed more frequently since CV$_{RED}$ incurs fewer violations than the attested CVC$_{RED}$ in DEP-OO-C, which outranks FTMIN, as illustrated in (54).

(54) *Constraints under the cut-off fail to rule out CV$_{RED}$ for CVC-BASE*

CVC-BASE	DEP-OO-C	FTMIN	DEP-OO-V	ANCHOR-IR$_\sigma$
💣$_1$ a. CV$_{RED}$	*	*	*	*
☞$_2$ b. CVC$_{RED}$	**		*	

Why is CV$_{RED}$ a possible variant for CV.C-BASE but not for CVC-BASE? The reason is also due to prosodic faithfulness: the matching between the syllable structure of the reduplicant and the underlying structure of the base is faithful in the former but not in the later because a CV$_{RED}$ corresponding to a CVC-BASE fails to copy an underlying coda. A generalization can thus be made: an undersized reduplicant is possible in Kavalan reduplication only if it can improve the IR matching of the syllable structure. In other words, a reduplicant cannot be undersized and at the same time not match the underlying syllable structure of the base. This can be captured by the conjunction of the two constraints FTMIN and ANCHOR-IR$_\sigma$. (56) illustrates that the placement of the conjoint constraint above the cut-off correctly rules out CV$_{RED}$ corresponding to CVC-BASE, but not to CV.C-BASE.

(55) [FTMIN & ANCHOR-IR$_σ$]$_{RED}$
A reduplicant cannot violate both FTMIN and ANCHOR-IR$_σ$.[18]

(56) *[FTMIN & ANCHOR-IR$_σ$]$_{RED}$ above the cut-off helps predict CV$_{RED}$ as a possible variant for CV.C-BASE but not for CVC-BASE*

a.

CVC-BASE	[FTMIN & ANCHOR-IR$_σ$]$_{RED}$	DEP-OO-C	FTMIN	DEP-OO-V	ANCHOR -IR$_σ$
☞$_1$ a. CVC$_{RED}$		**		*	
b. CV$_{RED}$	*!	*	*	*	*

b.

CV.C-BASE	[FTMIN & ANCHOR-IR$_σ$]$_{RED}$	DEP-OO-C	FTMIN	DEP-OO-V	ANCHOR -IR$_σ$
☞$_1$ a. CV$_{RED}$		*	*	*	
☞$_2$ b. CVC$_{RED}$		**		*	*

3.2.2.2 CCV-BASE

[FTMIN & ANCHOR-IR$_σ$]$_{RED}$ also helps to predict the post-vocalic C.CV$_{RED}$ yielded by CCV-BASE. Recall that [FTMIN & ANCHOR-IR$_σ$]$_{RED}$ requires that a reduplicant cannot be undersized and at the same time lack IR matching of the syllable structure. For CCV-BASE, any reduplicant candidates corresponding to CCV-BASE cannot surface without violating ANCHOR-IR$_σ$. That is because the only way for a reduplicant corresponding to CCV-BASE to achieve IR matching of syllable structure is to form an onset cluster; nonetheless, *COMP-M is inviolable in the language (e.g. /ʁa.-RED~**tβ**a.ʁi/ 'reddish' [L&T:455] > **ʁa.*-[(***tβa.***)]~*tβa.ʁiʔ*). Therefore, although FTMIN is ranked below the cut-off and, thus, it alone is unable to rule out undersized reduplicants such as Cɨ̯ (e.g. **ʁa.*-[(*tɨ̯*)]~*t.βa.ʁiʔ*) or a bare C that is syllabified as the coda of the preceding syllable (e.g. **ʁa*-[(*t*).]~*tɨ̯.βa.ʁiʔ*), undersized reduplicants corresponding to CCV-BASE (which must violate ANCHOR-IR$_σ$) necessarily lead to the violation of [FTMIN & ANCHOR-IR$_σ$]$_{RED}$ and will be ruled out before reaching the cut-off, as illustrated in (57). (57) also shows that DEP-IO/IR must rank above the cut-off to rule out Cɨ̯.CV$_{RED}$ (57b) in post-vocalic position.

[18] Under the theory of constraint conjunction (Smolensky 1993), only when both subparts of the constraint are violated will the conjoint constraint be violated.

(57) [*FtMin & Anchor-IR$_\sigma$*]$_{RED}$ *helps rule out post-vocalic undersized reduplicants that correspond to CCV-base*

/ʁa.-RED~tβa.ʁi/ ‘reddish’ [L&T:455] RO: tɨ̯.βa.ʁiʔ ‘red’	*Comp-M	FtMax	[FtMin & Anchor-IR$_\sigma$]$_{RED}$	Dep-IO/IR	Dep-OO-C	FtMin	Dep-OO-V	Anchor-IR$_\sigma$
☞1 a. ʁa-[(t$_\mu$.βa$_\mu$)]~t.βa.ʁiʔ					ʁ, t, β		a, a	*
b. ʁa.-[(tɨ̯$_\mu$.βa$_\mu$)]~t.βa.ʁiʔ				*!	ʁ, t, β		a, ɨ̯, a	*
c. ʁa-[(t$_\mu$.)]~tɨ̯.βa.ʁiʔ			*!	*	ʁ, t	*	a	*
d. ʁa.-[(tɨ̯$_\mu$.)]~tɨ̯.βa.ʁiʔ			*!	**	ʁ, t	*	a, ɨ̯	*
e. ʁa-[(t$_\mu$.βa$_\mu$ʁ$_\mu$.)]~tɨ̯.βa.ʁiʔ		*!		*	ʁ, t, β, ʁ		a, a	*
f. ʁa.-[(tβa$_\mu$)]~t.βa.ʁiʔ	*!		*		ʁ, t, β	*	a, a	*

For the same reason, [FtMin & Anchor-IR$_\sigma$]$_{RED}$ can help rule out Cɨ̯ corresponding to CCV-base in non-postvocalic position (ref. (58c)). But the function of the conjoint constraint is neutralized and unable to rule out Cɨ̯C in the same position because Cɨ̯C is bimoraic and, thus, does not violate FtMin. But Cɨ̯C violates *Ft/ɨ̯ because it is headed by *ɨ̯*. Therefore, to rule out Cɨ̯C as one of the reduplicant variants, *Ft/ɨ̯ must be placed above the cut-off, as illustrated in (58).

(58) **Ft*/ɨ̯ must rank above the cut-off*

/RED~q-zu.sa/ 'two or three days' [my field note] RO: qɨ̯zusaʔ 'two days'	[FtMin & Anchor-IR$_\sigma$]$_{RED}$	*Ft/ɨ̯	Dep-OO-C	FtMin	Dep-OO-V	Anchor-IR$_\sigma$
☞$_1$ a. [(qɨ̯$_\mu$.zu$_\mu$.)]~ qɨ̯.zu.saʔ			q, z		ɨ̯, u	*
b. [(qɨ̯$_\mu$z$_\mu$.)]~qɨ̯.zu.saʔ		*!	q, z		ɨ̯	*
c. [(qɨ̯$_\mu$.)]~qɨ̯.zu.saʔ	*!	*	q	*	ɨ̯	*

3.3 Reduplication in vowel initial bases

This section discusses reduplication in vowel initial bases, which includes V.CV-BASE and VC-BASE.

3.3.1 V.CV-BASE

In terms of syllable structure, V.CV-BASE is actually quite similar to CV.C-BASE in that both bases have the first syllable ends with a vowel, followed by the onset of a following syllable. They only differ in that V.CV-BASE starts without a phonemic onset but CV.C-BASE starts with one. Therefore, since CV.C-BASE yields three reduplicant variants: CV$_{RED}$, CVC$_{RED}$, and CV.CV$_{RED}$, V.CV-BASE is expected to yield reduplicant variants that take the shape of V, VC, and V.CV as well. However, V.CV-BASE only yields V.CV$_{RED}$. So why can't V.CV-BASE yield reduplicant variants that take the shape of V and VC?

Consider VC reduplicant first. Though VC reduplicant yielded by V.CV-BASE will be pretty much like the CVC$_{RED}$ yielded by CV.C-BASE (since both reduplicants end with a consonant), they differ in whether, when attaching to the base, the ending consonant will take the role as a coda (of a preceding vowel) or as an onset (of a following vowel). The ending consonant of CV**C**$_{RED}$ will be syllabified as a coda when prefixing to CV.C-BASE simply because Kavalan does not tolerate complex syllable margins (cf. (59a)). On the other hand, the ending consonant of a VC reduplicant will be syllabified as an onset when prefixing to V.CV-BASE because the base is without a phonemic onset and onset is universally preferred to coda (cf. (59b)). In other words, while CVC_{RED} is bimoraic, $V.C_{RED}$ is monomoraic.

(59) *The ending consonant of a VC reduplicant will be syllabified as an onset*

	Base	*Reduplicant shape*	*Syllabification of the reduplicated form*
a.	CV.C-BASE	CVC	<u>$CV_{\mu}C_{\mu}$.</u>~CV.C-BASE
b.	V.CV-BASE	VC	<u>V_{μ}.C</u>~V.CV-BASE

The monomoraic status of $V.C_{RED}$ is not fatal because FTMIN is ranked below the cut-off. However, since $V.C_{RED}$ only copies part of the second syllable of V.CV-BASE, it also violates ANCHOR-IR_{σ}. The combined violations of FTMIN and ANCHOR-IR_{σ} become fatal since it leads to the violation of [FTMIN & ANCHOR-$IR_{\sigma}]_{RED}$, which is ranked above the cut-off. Therefore, $V.C_{RED}$ is ruled out before reaching the cut-off and cannot be a reduplicant variant for V.CV-BASE, as illustrated in (60).

(60) *[FTMIN & ANCHOR-$IR_{\sigma}]_{RED}$ rules out $V.C_{RED}$ as a variant for V.CV-BASE*

V.CV-BASE	[FTMIN & ANCHOR-$IR_{\sigma}]_{RED}$
☞$_1$ a. <u>$V_{\mu}.CV_{\mu}$.</u>~V.CV-BASE	
b. <u>V_{μ}.C</u>~V.CV-BASE	*!

But the current constraint ranking cannot rule out V_{RED} as a reduplicant variant of V.CV-BASE. That is because though V_{RED} is monomoraic just like $V.C_{RED}$, it copies the entire first syllable of the base and no more, satisfying ANCHOR-IR_{σ}. It actually violates none of the constraints ranked above the cut-off and has to become a variant. It is proposed here that V_{RED} is bad for V.CV-BASE because the affixation of V_{RED} to V.CV-BASE will result in identical vocalic segments across morpheme boundary (e.g. *m-utiʔ* 'to vomit' > **m-**<u>u</u>~u**tiʔ* 'to keep vomiting' [L9:133]), violating OCP(VOC), which is assumed to rank above the cut-off.

(61) OCP(VOC)
Vocalic sequences of identical place features are disallowed. (e.g. *uu, *uw).[19]

[19] The example 'to keep catching birds' transcribed as *m-ʁi-aɬa-aɬam* in Lee (2009:133) might seem to be a counterexample to the constraint. However, since phonetic glottal stops before vowels are often left unmarked in Lee, 'to keep catching birds' could also be *m-ʁi-ʔaɬa-ʔaɬam* phonetically and conforms to OCP(VOC). The transcription of the phrase given in Li & Tsuchida (2006:24) is actually *m-ʁi-ʔaɬ-ʔaɬam*, which is with glottal stops and without *a* in adjacent position. Unfortunately, I failed to elicit the data from my Kavalan consultant.

(59) illustrates [FTMIN & ANCHOR-IR$_{\sigma}$]$_{RED}$ and OCP(VOC) ranked above the cut-off successfully predict V.CV$_{RED}$ as the only reduplicant for V.CV-BASE.

(62) *[FTMIN & ANCHOR-IR$_{\sigma}$]$_{RED}$ and OCP(VOC) help rule out V.C$_{RED}$ and V$_{RED}$ for V.CV-BASE*

/m-RED~u.ti/ 'to keep vomiting'	OCP(VOC)	[FTMIN & ANCHOR-IR$_{\sigma}$]$_{RED}$
☞$_{1}$ a. m-$\underline{u_{\mu}.ti_{\mu}}$~w.tiʔ		
b. m-$\underline{u_{\mu}.t}$~u.tiʔ		*!
c. m-$\underline{u_{\mu}.}$~u.tiʔ	*!	

3.3.2 VC-BASE

This section examines how VC$_{RED}$ is predicted for VC-BASE. In §3.3.1, it is shown that one of the reasons V.C$_{RED}$ cannot be a possible variant for V.CV-BASE is that when it is prefixed to a V.CV-BASE, the ending consonant in the reduplicant is syllabified as the onset of the following syllable, leaving the reduplicant monomoraic and undersized. However, the ending consonant of a VC reduplicant yielded by VC-BASE is always syllabified as a coda (i.e. VC.$_{RED}$ but not V.C$_{RED}$), as illustrated in (63) below, which is repeated from (12b) but added with syllable boundary.

(63) *VC reduplicant is syllabified as VC.$_{RED}$ for VC-BASE*

i. ʔiβ.ŋaw 'spider web' > ʔ$\underline{\text{iβ.}}$~iβ.ŋa.w-an 'a place full of spider webs' [my field note]
ii. ʔiʁ.βəŋ 'deep' > ʔ$\underline{\text{iʁ.}}$~iʁ.βəŋ 'very deep' [my field note]
iii. ʔuŋ.ɾaj 'pineapple' > ʔ$\underline{\text{uŋ.}}$~uŋ.ɾa.j-an 'a place full of pineapples' [my field note]

The fact the VC reduplicant corresponding to VC-BASE is syllabified as VC.$_{RED}$ rather than V.C$_{RED}$ ensures the bimoraic status of the reduplicant, making it free from the violation of [FTMIN & ANCHOR-IR$_{\sigma}$]$_{RED}$ (e.g. *ʔ$\underline{\text{i}_{\mu}\text{β}_{\mu}\text{.}}$~iβ.ŋaw-an* 'a place full of spider webs'). (64) illustrates how the current constraint ranking is sufficient to predict VC.$_{RED}$ (64a) as the only reduplicant for VC-BASE and how V.C$_{RED}$ (64b) is ruled out as a possible candidate.

(64) *VC$_{RED}$ is correctly predicted as the only reduplicant for VC-BASE*

/RED~iβ.ŋaw/ 'a place full of spider webs'	FTMAX	OCP(VOC)	[FTMIN & ANCHOR-IR$_{\sigma}$]$_{RED}$
☞$_{1}$ a. ʔ[($\underline{\text{i}_{\mu}\text{β}_{\mu}\text{.}}$)]~i$_{\underline{\mu}}$β$_{\underline{\mu}}$.ŋa$_{\underline{\mu}}w_{\underline{\mu}}$			
b. ʔ[($\underline{\text{i}_{\mu}\text{.β}}$)]~i$_{\underline{\mu}}$β$_{\underline{\mu}}$.ŋa$_{\underline{\mu}}w_{\underline{\mu}}$			*!
c. ʔ[($\underline{\text{i}_{\mu}\text{.}}$)]~i$_{\underline{\mu}}$β$_{\underline{\mu}}$.ŋa$_{\underline{\mu}}w_{\underline{\mu}}$		*!	*
d. ʔ[($\underline{\text{i}_{\mu}\text{β}_{\mu}\text{.ŋa}_{\mu}\text{.}}$)]~i$_{\underline{\mu}}$β$_{\underline{\mu}}$.ŋa$_{\underline{\mu}}w_{\underline{\mu}}$	*!		

However, before ending the analysis for VC-BASE, we need to go back to the analysis for V.CV-BASE. In (62), VC.$_{RED}$ is not considered in the candidate pull for V.CV-BASE (i.e. **m-<u>u$_{\mu}$t$_{\mu}$.</u>~u.tiʔ*). Nonetheless, since the VC reduplicant corresponding to VC-BASE is syllabified as VC.$_{RED}$, the VC reduplicant corresponding to V.CV-BASE should be able to be syllabified as VC.$_{RED}$, too, escaping the violation of [FTMIN & ANCHOR-IR$_{\sigma}$]$_{RED}$. (65) shows that not only can [FTMIN & ANCHOR-IR$_{\sigma}$]$_{RED}$ not rule out VC.$_{RED}$ as a possible reduplicant for V.CV-BASE, but that DEP-OO-V under the cut-off will also wrongly predict VC.$_{RED}$ as a more commonly observed variant than the attested V.CV$_{RED}$.

(65) *The current constraint ranking fails to rule out VC.$_{RED}$ as a variant for V.CV-BASE*

/m-RED~uti/ 'to keep vomiting' RO: mutiʔ 'to vomit'	[FTMIN & ANCHOR-IR$_{\sigma}$]$_{RED}$	DEP-OO-C	FTMIN	DEP-OO-V
☞$_2$ a. m-<u>u$_{\mu}$.ti$_{\mu}$</u>~w.tiʔ		t		u, i!
💣$_1$ b. m-<u>u$_{\mu}$t$_{\mu}$.</u>~u.tiʔ		t		u

It is proposed that the reason a VC reduplicant corresponding to VC-BASE can be syllabified as VC.$_{RED}$ while that corresponding to V.CV-BASE cannot is due to the σ-ROLE-BR constraint, which requires corresponding segments in the reduplicant and the base to have identical syllable role. The VC-BASE~VC.$_{RED}$ correspondence satisfies σ-ROLE-BR because the coda segment in the base nicely corresponds to a coda segment in the reduplicant. Nonetheless, V.CV-BASE cannot yield VC.$_{RED}$ because in that case the onset segment in the base will correspond to a coda segment in the reduplicant, violating σ-ROLE-BR. Compare (67a) with (67b) below.

(66) σ-ROLE-BR
Corresponding segments in the reduplicant and the base must have identical syllable roles (cf. Rose & Walker 2004, Yu 2005, and others).

(67) *σ-ROLE-BR predicts VC.$_{RED}$ can be yielded by VC.-BASE but not V.CV-BASE*

	Base	*Reduplicant*	*Syllable role in BR correspondence*	σ-ROLE-BR
a.	VC.-BASE	VC.$_{RED}$	<u>V**C**.</u> ~V**C**.-BASE (e.g. *ʔ<u>ur.</u>~ur.sa.p-an* 'a place full of chicken fleas' [my field note])	✓
b.	V.CV-BASE	VC.$_{RED}$	<u>V**C**.</u> ~V.**C**V-BASE (e.g. **m-<u>at.</u>~a.**t**iw* 'to go often', cf. attested: *m-<u>a.**tj**</u>~a.**t**iw*)	✗

However, in Kavalan violations of σ-ROLE-BR can be observed in the CVC.$_{RED}$~CV.C-BASE correspondence and the Cɨ̯.VC.$_{RED}$~CCV-BASE correspondences, as illustrated in (68).

(68) *σ-ROLE-BR is violated in CVC.$_{RED}$~CV.C-BASE and Cɨ̯.VC.$_{RED}$~CCV-BASE*

	Base	*Reduplicant*	*Syllable role in BR correspondence*	σ-ROLE-BR
a.	CV.C-BASE	CVC.$_{RED}$	CV**C**. ~CV.**C**-BASE (e.g. s-ɨ̯.m-aŋ.~sa.ŋiʔ 'to keep making' [L&T:370])	×
b.	CCV-BASE	Cɨ̯.CV$_{RED}$	**Cɨ̯**.CV ~**C**.CV-BASE (e.g. ***mɨ̯***.*ɾi*~**m**.*ɾi.mək* 'taciturn' cf. /*mɾimək*/ [L&T:325])	×

Careful examination of the attested correspondences in (68) and the unattested correspondence in (67b) reveals that σ-ROLE-BR is violable in Kavalan, unless the violation of the constraint will result in an onsetless syllable, just as in (67b). Thus, the conjunction of σ-ROLE-BR and ONSET (i.e. [σ-ROLE-BR & ONSET]$_{WD}$) can help rule out (67b) but not (67a) nor (68a-b) because (67a) violates ONSET but not σ-ROLE-BR and (68a) and (68b) violates σ-ROLE-BR but not ONSET.[20] [σ-ROLE-BR & ONSET]$_{WD}$ is assumed to be undominated in Kavalan.

(69) [σ-ROLE-BR & ONSET]$_{WD}$
A word cannot violate both σ-ROLE-BR and ONSET.

3.4 Shape invariance in Kavalan reduplication

In §3.1-§3.3 we have provided an account to Kavalan reduplication. Universally reduplicants tend to have an invariant shape that has no one-to-one relation with a prosodic constituent in the base. Thus, languages seldom copy, for example, the first

[20] Notice that V.CV$_{RED}$ corresponding to V.CV-BASE will also result in a violation of σ-ROLE-BR. As mentioned in fn.9, gliding is observed in the reduplication of V.CV-BASE, changing the initial V in the base to a glide. That is, V.C**V**~**V.**CV-BASE → V.C**V**~**G**.CV-BASE (ref. (12a)). Thus, the initial V of the reduplicant, which takes the role of nucleus, will correspond to the glide at the beginning of the base, which takes the role of coda, violating σ-ROLE-BR. Though the V.CV$_{RED}$~V.CV-BASE correspondence violates σ-ROLE-BR, it will not violate [σ-ROLE-BR & ONSET]$_{WD}$ since the violation of σ-ROLE-BR in the V.CV$_{RED}$~V.CV-BASE correspondence is actually to prevent an ONSET violation, not to create one.

syllable exactly regardless of whether it has a coda or not. However, in Kavalan, the first syllable of the base plays an important role in conditioning the reduplicant shape. With the exception of V.CV-BASE, the reduplicant always copies the first syllable of the base regardless of whether it has a coda or not, as summarized in (70), column A.[21] In this respect, shape invariance seems to play no role in Kavalan. On the other hand, Kavalan is trying to maintain a fixed reduplicant size. It is argued that the default size of the Kavalan reduplicant is bimoraic. Even though a reduplicant may be undersized when the base initial syllable is monomoraic, for those bases that yield an undersized monomoraic reduplicant [i.e. (70c) CV.C-BASE and (70d) CV.V-BASE], the undersized reduplicant always has bimoraic variants (ref. (70), column B), suggesting that Kavalan is also trying to maintain an invariant size.

(70) *Reduplication of base initial syllable vs. shape invariance*

	A		B	
	reduplication of base initial syllable?		*base initial syllable undersized*?	
a. CVC-BASE	$\xrightarrow{\text{yes}}$	$\mathrm{CVC}_{\mathrm{RED}}$	$\xrightarrow{\text{no}}$	$\mathrm{CVC}_{\mathrm{RED}}$
b. CCV-BASE	$\xrightarrow{\text{yes}}$	$\mathbf{CCV}_{\mathrm{RED}}$ (post-vocalic) **Cɨ̯CV**$_{\mathrm{RED}}$ (elsewhere)	$\xrightarrow{\text{no}}$	$\mathrm{CCV}_{\mathrm{RED}}$ (post-vocalic) Cɨ̯CV$_{\mathrm{RED}}$ (elsewhere)
c. CV.C-BASE	$\xrightarrow{\text{yes}}$	$\mathrm{CV}_{\mathrm{RED}}$	$\xrightarrow{\text{yes}}$	$\mathrm{CV}_{\mathrm{RED}}$ ~ $\mathbf{CVC}_{\mathbf{RED}}$ ~ $\mathbf{CVCV}_{\mathrm{RED}}$ (49%) (39%) (11%)
d. CV.V-BASE	$\xrightarrow{\text{yes}}$	$\mathrm{CV}_{\mathrm{RED}}$	$\xrightarrow{\text{yes}}$	$\mathrm{CV}_{\mathrm{RED}}$ ~ $\mathbf{CVV}_{\mathbf{RED}}$ (29%) (71%)
e. VC-BASE	$\xrightarrow{\text{yes}}$	$\mathrm{VC}_{\mathrm{RED}}$	$\xrightarrow{\text{no}}$	$\mathrm{VC}_{\mathrm{RED}}$
f. V.CV-BASE	$\xrightarrow{\text{no}}$ [blocked by OCP(VOC)]	n/a	$\xrightarrow{\text{n/a}}$	$\mathrm{V.CV}_{\mathrm{RED}}$

The framework of Optimality Theory, and ROE (Coetzee 2006) in particular, gives us the flexibility to capture the conflicting forces in Kavalan. FTMIN, which encourages shape invariance and ANCHOR-IR$_{\sigma}$, which encourages the reduplicant to copy the underlying prosodic unit from the base, are ranked below the cut-off. Therefore, exact copying of the base initial syllable and copying of an invariant prosodic unit independent of the base both surface as variants of a base. However, not every base produces variants; the

[21] As mentioned in §3.3.1, the reduplicant of V.CV-BASE cannot be the copy of the first syllable because it would violate OCP(VOC).

conjunction of FTMIN and ANCHOR-IR$_σ$ (i.e. [FTMIN & ANCHOR-IR$_σ$]$_{RED}$) ranked above the cut-off serves to capture the fact that only when exact copying of base initial syllable would generate an undersized reduplicant will variation occur.

A summary of the constraint ranking developed in this section to account for the various reduplicant shapes is given in (71).

(71) *Constraint hierarchy to predict the reduplicant shapes*
HEADNESS, RED-PRWD-L, RED-PRWD-R, *FT/ɨ̥, FTMAX,
[FTMIN & ANCHOR-IR$_σ$]$_{RED}$, LINEARITY-IR, OCP(VOC), [σ-ROLE-BR & ONSET]$_{WD}$,
\>> DEP-IO/IR
---cut-off---
DEP-OO-C
\>> FTMIN
\>> DEP-OO-V, ANCHOR-IR$_σ$
\>> MAX-BR

4. The apparent counterexamples

As mentioned in §2, there are two types of examples that do not conform to the Base~Reduplicant correlation (cf. (2) and (16)), one involves CVC-BASE that is shown to yield CV$_{RED}$ rather than the expected CVC$_{RED}$ and the other involves V.CV-BASE that is shown to yield V.C$_{RED}$ rather than the expected V.CV$_{RED}$. The two types of examples are considered in §4.1 and §4.2, respectively.

4.1 Examples involving CVC-BASE

Careful examination of the examples involving CVC-BASE in (13) reveals that they are of two types: those where the first syllable of the base ends with a geminate (13a-d) and those where the root is monosyllabic (13e-f).

4.1.1 Geminate ending CVC-BASE

Examples (13a-d) are repeated below for ease of reference.[22] (Geminate clusters are in boldface here below).

[22] A reviewer points out that some of the words with internal geminates may have non-geminate variants. For instance, "to sneeze" may be pronounced as *ßasiŋ* or *ßassiŋ*. Base variants without internal geminates belong to CV.C-BASE (rather than CVC-BASE) and naturally yield CV$_{RED}$ (cf. §3.2.1.1).

(72) *CVC-BASE ending with a geminate unexpectedly yields CV_{RED}*

a. ɾuβi**ŋŋ**aw	'become dizzy'	>	ɾu~<u>βi</u>~βi**ŋŋ**aw	'very dizzy'	[L&T:87]
b. na**pp**aw-an	'spouse'	>	<u>na</u>~na**pp**aw	'to marry'	[L&T:197]
c. qa-ɾu**ss**iq	'one'	>	qa-<u>ɾu</u>~ɾu**ss**iq	'one for each person'	[L&T:226]
d. βa**ss**iŋ	'to sneeze'	>	<u>βa</u>~βa**ss**iŋ	'to keep sneezing'	[L&T:82]

Why does CVC-BASE with the first syllable ending with a geminate yield CV_{RED} rather than CVC_{RED}? This actually has to do with the inherent nature of geminates. One of the most remarkable features of geminates is integrity (Kenstowicz & Pyle 1973, Guerssel 1978, Hayes 1986, Schein & Steriade 1986, among others). Geminates are typically known to be inseparable (not allowing insertion of an intervening segment) and inalterable (not tolerating modifications that would affect only one part of the segment). Thus, it is common to see rules of metathesis, copying, epenthesis, etc., being blocked if their application should result in the separation of a geminate cluster. Geminate integrity provides an answer as to why the reduplicant of such a CVC-BASE cannot be the normal CVC_{RED}; if the reduplicant is of the shape CVC, only part of the geminate cluster would be copied, violating the GEM-INTEGRITY constraint, which prohibits the reduplicant from copying part of the geminate cluster.

There are two possible ways to avoid violating GEM-INTEGRITY: one is to copy neither of the segments in the geminate cluster (yielding CV_{RED}), the other is to copy both of the geminate segments (yielding CV**CC**$_{RED}$). As shown in §3.2.2.1, CV_{RED} corresponding to CVC-BASE would violate [FTMIN & ANCHOR-IR_{σ}]$_{RED}$. On the other hand, CV**CC**$_{RED}$ involves complex clusters at syllable edges, violating *COMP-M. As the attested reduplicant is CV_{RED}, *COMP-M and GEM-INTEGRITY must dominate [FTMIN & ANCHOR-IR_{σ}]$_{RED}$. (73) illustrates how the constraint ranking predicts CV_{RED} for a geminate ending CVC-BASE.

(73) *GEM-INTEGRITY must outrank [FTMIN & ANCHOR-IR_{σ}]$_{RED}$*

/RED~βa**ss**iŋ/ 'to keep sneezing'	*COMP-M	GEM-INTEGRITY	[FTMIN & ANCHOR-IR_{σ}]$_{RED}$
☞$_1$ a. <u>βa</u>~βa**ss**iŋ			*
b. <u>βas</u>~βa**ss**iŋ		*!	
c. <u>βass</u>~βa**ss**iŋ	*!		

4.1.2 Monosyllabic CVC-BASE

The other type of CVC-BASE that yields CV_{RED} rather than CVC_{RED} is when the base is monosyllabic. Examples (13e-f) are repeated below for ease of reference:

(74) *Monosyllabic CVC-BASE unexpectedly yields CV_{RED}*

a.	ma-βuɫ	'fist fight'	>	ma-βu~βuɫ	'fight together'	[L&T:152]
b.	sum	'urine'	>	su-su~sum	'smell of urine'	[L10:103]

For monosyllabic CVC-BASE, if the reduplicant has the form of CVC, the reduplicant and the base would be identical. Though it is natural for reduplicant and base to be identical, in some languages there is a tendency to prevent total identity between the base and the reduplicant. For instance, Kenstowicz (1985) and Yip (1995) report examples from Javanese habitual repetitive reduplication in which the reduplicant and the base are prevented from containing identical elements, as illustrated in (75) cited from Hicks-Kennard (2004:321).

(75) *OCP effect in Javanese habitual repetitive reduplication* (*Kenstowicz 1985*, *Yip 1995*)

a.	eli	ela~eli	'remember'	*eli~eli
b.	tuku	tuka~tuku	'buy'	*tuku~tuku
c.	ele	ela~ele	'bad'	*ele~ele
d.	bul	bal~bul	'puff'	*bul~bul

Yip (1995:23) proposes the OCP(STEM) constraint, which prohibits total identity between the reduplicant and the base, to account for the effect in Javanese habitual reduplication. The OCP(STEM) constraint is adopted and ranked above $[\text{FtMin} \ \& \ \text{Anchor-IR}_{\sigma}]_{\text{RED}}$, which favors CVC_{RED}. $\|\text{OCP(STEM)} >> [\text{FtMin} \ \& \ \text{Anchor-IR}_{\sigma}]_{\text{RED}}\|$ correctly predicts the lack of copying of the coda from a monosyllabic CVC-BASE, as illustrated in (76).[23]

(76) *$\|\text{OCP(STEM)} >> [\text{FtMin} \ \& \ \text{Anchor-IR}_{\sigma}]_{\text{RED}}\|$ predicts CV_{RED} for monosyllabic CVC-BASE*

/ma-RED~βuɫ/ 'fight together'	OCP(STEM)	$[\text{FtMin} \ \& \ \text{Anchor-IR}_{\sigma}]_{\text{RED}}$
☞1 a. ma-βu~βuɫ		*
b. ma-βuɫ~βuɫ	*!	

[23] Lee (2009:fn3) remarks that though roots beginning with a CV.V string tend to yield $CV.V_{RED}$, the reduplicant of trisegmental CV.V words can only be CV- (e.g. *mai* 'none' > *sia-ma~mai* 'become fewer and fewer/less and less', **sia-mai~mai*). But there is a counterexample: *kia* 'little' > *kia~kia* 'little for each' [my field note] (cf. also Li & Tsuchida 2006:122). It could be that the *ai* sequence in *mai* undergoes gliding and changes to *aj* to avoid vowel cluster (i.e. *sia-ma~maj*). For the *ia* sequence in *kia*, no matter whether gliding takes place or not (i.e. *kja* vs. *kia*), the vowel *a* is followed by a glottal stop in the output (as predicted by ||SWP >> DEP-IO||) (e.g. *kia~kiaʔ*). If so, OCP(STEM) can help explain why **sia-maj~maj* is ill-formed while *kia~kiaʔ* is well-formed because the former, but not the latter, contains an identical element in the reduplicant and the base.

The OCP(STEM) constraint also provides explanations to the examples in (77) in which a process of gemination occurs changing the coda of CVC_{RED} to be the same as the onset of the following syllable.

(77) *More OCP effect in Kavalan*

a.	mɨ̯-ta**ɬ**	‘to defecate’	>	mɨ̯-<u>ta**t**</u>~**t**aɬ	‘to have diarrhea’	[L&T:439]
b.	mɨ̯-za**m**	‘to catch up’	>	mɨ̯-<u>za**z**</u>~**z**am	‘to catch up’	[L&T:517]
c.	ʁa**q**	‘alcohol’	>	su-<u>ʁa**ʁ**</u>~**ʁ**aq	‘smell of alcohol’	[L9:144]
d.	mɨ̯-ŋa**ʁ**	‘to open one’s mouth widely’	>	mɨ̯-<u>ŋa**ŋ**</u>~**ŋ**aʁ	‘to open one’s mouth repetitively’	[L&T:213]

In Kavalan, such a geminating process occurs sporadically when the base is disyllabic or longer, as illustrated in (78). But when the base is the monosyllabic CVC-BASE and yields CVC_{RED}, gemination always occurs (cf. (77)). In other words, no matter whether the reduplicant of a monosyllabic CVC-BASE surfaces as CV_{RED} (cf. (74)) or as CVC_{RED} (cf. (77)), the segmental information in the reduplicant and the base is never the same, satisfying OCP(STEM).

(78) *Gemination occurs sporadically when the root is not monosyllabic*

	ai.	puɾan	‘to tie a knot at the end of string’	>	<u>pu**p**</u>~**p**uɾan	‘to tie a knot at the end of a string repetitively’	[L&T:253]
But:							
	aii.	βaɾiʔ	‘wind’	>	<u>βaɾ</u>~**β**aɾi-an *<u>βa**β**</u>**~β**aɾi-an	‘a place where the wind is blowing all the time’	[L&T:79]
	aiii.	mɨ̯-quɾin	‘to roll without intention’	>	mɨ̯-<u>quɾ</u>~**q**uɾin *mɨ̯-<u>qu**q**</u>~**q**uɾin	‘to keep rolling’	[my field note]
	bi.	sa-qu**ŋ**aʔ	‘to lie’	>	sa-<u>qu**q**</u>~**q**uŋaʔ	‘liar’	[L&T:304]
But:							
	bii.	qiʁəβ	‘smoke’	>	su-<u>qi**ʁ**</u>~**q**iʁəβ *su-<u>qi**q**</u>**~q**iʁəβ	‘smell of smoke’	[L9:143]
	biii.	si**ŋ**uit	‘to blow one’s nose’	>	<u>si**ŋ**</u>~**s**iŋuit *si<u>**s**</u>~**s**iŋuit	‘to keep blowing one’s nose’	[L&T:217]

There are two things worth noting about the OCP(stem) constraint. First, the constraint would predict that there is no total reduplication in Kavalan. Examples of reduplication that involve total copying are indeed rare in Kavalan. There are only two examples available in the literature.[24]

[24] Notice that words involving lexicalized reduplication (e.g. *kuskus* ‘scratch’) cannot be considered as total reduplication since the non-reduplicated parts cannot stand alone. In

(79) *Reduplication that involve total copying is rare in Kavalan*

ʔəɬan	'sky, day'	>	<u>ʔəɬan</u>~ʔəɬan	'every day'	[C:60]
tasaw	'year'	>	<u>tasaw</u>~tasaw	'every year'	[C:60]

There are two possible explanations. The first is to limit the function of OCP(STEM) to monosyllabic bases. The other is to consider the two examples in (79) as compound forms (cf. Li & Tsuchida 2001).

Second, the OCP(STEM) constraint is in conflict with BR correspondence constraints. It must outrank MAX-BR and IDENT-BR to predict the lack of coda copying and the gemination process found in a monosyllabic CVC-BASE. Furthermore, OCP(STEM) must be dominated by ANCHOR-BR-L (cf. (18)) and IDENT-BR-V, which requires BR correspondence of a vowel, to ensure the onset and the nucleus of the reduplicant will be identical to their base correspondents (e.g. /mɨ̯.-<u>RED</u>~taɬ/ 'to have diarrhea' → [mɨ̯.-<u>**tat**</u>.~taɬ], *[mɨ̯.-<u>**saɬ**</u>.~taɬ], *[mɨ̯.-<u>**tiɬ**</u>.~taɬ] [L&T:439]).

4.2 Examples involving V.CV-BASE

The data in (14), which are repeated below, contain examples of V.CV-BASE that unexpectedly yields V.C_{RED} rather than V.CV_{RED}.

(80) *V.CV-BASE unexpectedly yields V.C_{RED}*

a.	m-ipiɾ	'to listen'	>	m-<u>ip</u>~ipiɾ	'to keep listening'	[my field note]
b.	m-ataʁ	'dirty'	>	m-<u>at</u>~ataʁ	'very dirty'	[my field note]
c.	m-isis	'to carry'	>	m-<u>is</u>~isis	'to keep carrying'	[my field note]
d.	m-aɾaʔ	'to take'	>	ʔ<u>aɾ</u>~aɾa-n	'to keep taking'	[my field note]

As mentioned in §3.3.1, for V.CV-BASE, if the reduplicant has the form of V.C_{RED} the reduplicant will be monomoraic and copies only part of a base syllable, violating both FTMIN and ANCHOR-IR$_{\sigma}$ and be ruled out by [FTMIN & ANCHOR-IR$_{\sigma}$]$_{RED}$. Careful examination of data in (80), where V.CV-BASE unexpectedly yields VC_{RED}, however, reveals something interesting. All of the examples in (80) contain identical vowels in the first two syllables of the base. Therefore, if the reduplicant takes the form of V.CV,

addition, examples such as those listed below, which are generally transcribed as involving total reduplication, are not considered as such in the present study because the word-final glottal stop, which is generally left untranscribed in the literature, is not copied.

non-reduplicated form			*reduplicated form*		*generally transcribed as*	
wasuʔ	'dog'	>	su-<u>wasu</u>~wasuʔ	'having a smell of dogs'	su-<u>wasu</u>~wasu	[L9:133, L&T:22]
ɬamuʔ	'village'	>	<u>ɬamu</u>~ɬamuʔ	'every village'	<u>ɬamu</u>~ɬamu	[L7:266, L9:142]
kiaʔ	'little, few'	>	<u>kia</u>~kiaʔ	'little for each'	<u>kia</u>~kia	[my field note]

the reduplicated form will have vocalic elements of the same place standing in the morpheme boundary (e.g. *m-ipiɾ* 'to listen' > *m-**i**p**i**~**j**piɾ 'to keep listening'), violating the OCP(VOC) proposed above in (61). (81) illustrates how OCP(VOC), which is ranked above [FTMIN & ANCHOR-IR$_σ$]$_{RED}$, readily predicts V.C$_{RED}$ for V.CV-BASE that has identical vowel in the first two syllables.

(81) *||OCP(VOC) >> [FTMIN & ANCHOR-IR$_σ$]$_{RED}$|| predicts V.C$_{RED}$ for V.CV-BASE with identical vowel in the first two syllables*

/m-RED~i.piɾ 'to keep listening'	OCP(VOC)	[FTMIN & ANCHOR-IR$_σ$]$_{RED}$
☞$_1$ a. m-[(i$_μ$.p)]~**i.**piɾ		*
b. m-[(i$_μ$.**pi**$_μ$)]~**j.**piɾ	*!	

4.3 Interim summary

In sum, the examples in (13) and (14), which respectively involve a CVC-BASE yielding CV$_{RED}$ and a V.CV-BASE yielding V.C$_{RED}$ do not construct real counterexamples to the CVC-BASE~CVC$_{RED}$ and the V.CV-BASE~V.CV$_{RED}$ correlation but are triggered by other effects in the language. The former can be accounted for by introducing and ranking GEM-INTEGRITY and OCP(STEM) above [FTMIN & ANCHOR-IR$_σ$]$_{RED}$ and the latter falls naturally by the already existing ||OCP(VOC) >> [FTMIN & ANCHOR-IR$_σ$]$_{RED}$|| ranking. Number (82) summarizes the full constraint ranking concerning the treatment of Kavalan reduplication (including syllable structure constraints).

(82) *Full constraint hierarchy*

HEADNESS, RED-PRWD-L, RED-PRWD-R, *FT/i̯, FTMAX, ANCHOR-BR-L, OCP(VOC), [σ-ROLE-BR & ONSET]$_{WD}$, IDENT-BR-V, INITIAL-C, *COMP-M, SWP, *LONG-V, LINEARITY-IR,
>> OCP(STEM), GEM-INTEGRITY
>> [FTMIN & ANCHOR-IR$_σ$]$_{RED}$, IDENT-BR
>> DEP-IO/IR
--------------------------------------cut-off--------------------------------------
DEP-OO-C
>> FTMIN
>> DEP-OO-V, ANCHOR-IR$_σ$
>> MAX-BR

5. Conclusion

In this paper, we have provided an account of the reduplicative morpheme in Kavalan, which takes several distinct shapes depending on various properties of the base. Instead of considering Kavalan reduplication as involving two distinct reduplicant sizes (one monosyllabic and the other disyllabic), as suggested by the pioneering work of Lee (2009), the author of this paper argues instead that despite the several distinct shapes of the reduplicant, they share a default size of bimoraicity. In addition, this paper also argues that Kavalan reduplication is being torn between two contradictory forces. On the one hand, the reduplicant tries to copy an underlying prosodic unit from the base; on the other hand, it also tries to maintain an invariant shape which is bimoraic in size. The former force is captured by the prosodic faithfulness constraint ANCHOR-IR$_{\sigma}$, while the latter mainly by the size constraint FTMIN. The interaction of the two constraints results in variations in some bases and lack of variation in others. The framework of Optimality Theory and ROE gives us the flexibility to capture the conflicting forces in the language. The conjunction of the two constraints (i.e. [FTMIN & ANCHOR-IR$_{\sigma}$]$_{RED}$) above the cut-off predicts that no variation occurs when exact copying of the base initial syllable structure fulfills the bimoraic size requirement. This is the case in CVC-BASE~CVC$_{RED}$ and VC-BASE~VC$_{RED}$ correspondence. When the exact copying of the base initial syllable would produce an undersized reduplicant as in the case of CV.C-BASE and CV.V-BASE, variations of the reduplicant formed by copying the exact base initial syllable (i.e. CV$_{RED}$) and by copying an invariant (bimoraic) size (e.g. CV.C$_{RED}$) both occur. The relevant frequency of the variants is predicted by the relevant ranking of the constraints below the cut-off. For CCV-BASE, syllable structure constraints in the language such as *COMP-M make it impossible to copy the exact prosodic structure of the base initial syllable. But the copying of all segments in the base initial syllable (coupled with vowel insertion in non-post-vocalic position) would fulfill the requirement of bi-moracity. Unlike CV.C-BASE and CV.V-BASE, undersized CV$_{RED}$ and the perfect bimoraic CVC$_{RED}$ are never possible variants for CCV-BASE because they would be headed by an inserted weak vowel, violating the dominant *FT/ɨ̭ constraint in the language.

References

Alderete, John. 1996. Prosodic faithfulness in Cupeño. Manuscript. Amherst: University of Massachusetts.

Alderete, John, Jill N. Beckman, Laura Benua, Amalia Gnanadesikan, John J. McCarthy, and Suzanne Urbanczyk. 1999. Reduplication with fixed segmentism. *Linguistic Inquiry* 30.3:327-364.

Benua, Laura. 1997. *Transderivational Identity: Phonological Relations Between Words*. Amherst: University of Massachusetts dissertation.

Blust, Robert A. 1998. Ca- reduplication and proto-Austronesian grammar. *Oceanic Linguistics* 37.1:29-64.

Chang, Henry Y. 2000. *Gemalanyu Cankao Yufa* [*A Reference Grammar of Kavalan*]. Taipei: Yuan-Liou.

Chen, Matthew Y. 2000. *Tone Sandhi: Patterns Across Chinese Dialects.* Cambridge & New York: Cambridge University Press.

Coetzee, Andries W. 2006. Variation as accessing 'non-optimal' candidates. *Phonology* 23.3:337-385.

Crowhurst, Megan. 2004. Mora alignment. *Natural Language and Linguistic Theory* 22.1:127-177.

De Lacy, Paul. 2004. Markedness conflation in Optimality Theory. *Phonology* 21.2: 145-199.

Downing, Laura J. 2000. Satisfying minimality in Ndebele. *ZAS Papers in Linguistics* 19:23-39.

Everett, Daniel L. 1996. Prosodic levels and constraints in Banawá and Suruwahá. Available at Rutgers Optimality Archive (http://roa.rutgers.edu/, ROA#121-0496).

Everett, Daniel L. 2003. Iambic feet in Paumari and the theory of foot structure. *Linguistic Discovery* 2.1:22-44.

Ezard, Bryan. 1997. *A Grammar of Tawala: An Austronesian Language of the Milne Bay Area*, *Papua New Guinea*. Canberra: Pacific Linguistics.

Gafos, Diamandis. 1998. A-templatic reduplication. *Linguistic Inquiry* 29.3:515-527.

Gouskova, Maria. 2003. *Deriving Economy: Syncope in Optimality Theory*. Amherst: University of Massachusetts dissertation.

Gouskova, Maria. 2004. Minimal reduplication as a paradigm uniformity effect. *Proceedings of the 23rd West Coast Conference on Formal Linguistics* (*WCCFL-23*), ed. by Vineeta Chand, Ann Kelleher, Angelo J. Rodríguez & Benjamin Schmeiser, 265-278. Somerville: Cascadilla Press.

Guerssel, Mohamed. 1978. A condition on assimilation rules. *Linguistic Analysis* 4.3: 225-254.

Hayes, Bruce. 1986. Inalterability in CV phonology. *Language* 62.2:321-351.

Hendricks, Sean Q. 1999. *Reduplication without Template Constraints: A Study in Bare-Consonant Reduplication*. Tucson: University of Arizona dissertation.

Hicks-Kennard, Catherine. 2004. Copy but don't repeat: the conflict of dissimilation and reduplication in the Tawala durative. *Phonology* 21.3:303-323.

Hsiao, Yuchau E. 2000. Trisyllabic and quadrasyllabic Hakka tone sandhi: an Optimality Theory perspective. *Proceedings of the NCCU Teachers' Conference on Linguistics Research*, 101-124. Taipei: National Chengchi University.

Hsiao, Yuchau E. 2008. Yinping tone sandhi in two Hakka dialects. *Interfaces in Chinese Phonology: Festschrift in Honor of Matthew Y. Chen on his 70th Birthday*, ed. by Yuchau E. Hsiao, Hui-chuan Hsu, Lian-Hee Wee & Dah-an Ho, 79-97. Taipei: Institute of Linguistics, Academia Sinica.

Hsieh, Fuhui. 2007. *Language of Emotion and Thinking in Kavalan and Saisiyat.* Taipei: National Taiwan University dissertation.

Inkelas, Sharon. 1999. Exceptional stress-attracting suffixes in Turkish: representations versus the grammar. *The Prosody-Morphology Interface*, ed. by René Kager, Harry van der Hulst & Wim Zonneveld, 134-187. Cambridge & New York: Cambridge University Press.

Itô, Junko, Yoshihisa Kitagawa, and Armin Mester. 1996. Prosodic faithfulness and correspondence: evidence from a Japanese argot. *Journal of East Asian Linguistics* 5.3:217-294.

Jiang, Haowen. 2006. *Spatial Conceptualizations in Kavalan*. Taipei: National Taiwan University MA thesis.

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

Kennedy, Robert. 2005. The binarity effect in Kosraean reduplication. *Phonology* 22.2: 145-168.

Kenstowicz, Michael J. 1985. Multiple linking in Javanese. *Proceedings of the 16th Annual Meeting of the North East Linguistic Society* (*NELS 16*), ed. by Stephen Berman, Jae-Woong Choe & Joyce McDonough, 230-248. Amherst: GLSA.

Kenstowicz, Michael J. 1997. Quality-sensitive stress. *Rivista di Linguistica* 9.1:157-188.

Kenstowicz, Michael J., and Charles Pyle. 1973. On the phonological integrity of geminate clusters. *Issues in Phonological Theory*, ed. by Michael J. Kenstowicz & Charles W. Kisseberth, 27-43. The Hague: Mouton.

Kurisu, Kazutaka. 2001. *The Phonology of Morpheme Realization*. Santa Cruz: University of California dissertation.

Lee, Amy Pei-jung. 2007. *A Typological Study on Reduplication in Formosan Languages*. Colchester: University of Essex dissertation.

Lee, Amy Pei-jung. 2009. Kavalan reduplication. *Oceanic Linguistics* 48.1:130-147.

Lee, Amy Pei-jung. 2010. Reduplication and odor in four Formosan languages. *Language and Linguistics* 11.1:99-126.

Li, Paul Jen-kuei. 1982. Kavalan phonology: synchronic and diachronic. *GAVA: Studies in Austronesian Languages and Cultures Dedicated to Hans Kähler*, ed. by Rainer Carle, Martina Heinschke, Peter W. Pink, Christel Rost & Karen Stadtlander, 479-495. Berlin: Dietrich Reimer Verlag.

Li, Paul Jen-kuei. 1996. *Yilanxian Nandao Minzu yu Yuyan* [*The Austronesian Ethnic Groups and Languages of Yilan County*]. Yilan: Yilan County Government.

Li, Paul Jen-kuei, and Shigeru Tsuchida. 2001. *Pazih Dictionary.* Taipei: Institute of Linguistics, Academia Sinica.

Li, Paul Jen-kuei, and Shigeru Tsuchida. 2006. *Kavalan Dictionary.* Taipei: Institute of Linguistics, Academia Sinica.

Lin, Dong-yi. 2006. *The Language of Emotion in Kavalan*. Taipei: National Taiwan University MA thesis.

Lin, Hui-shan. 2010. Disyllabic verbal reduplication in Pazih—Leftward or rightward? *Language and Linguistics* 11.4:679-733.

Lin, Ju-en. 1996. *Tense and Aspect in Kavalan*. Hsinchu: National Tsing Hua University MA thesis.

McCarthy, John J. 1995. Extensions of faithfulness: Rotuman revisited. Manuscript. Amherst: University of Massachusetts. Available at: http://works.bepress.com/john_j_mccarthy/59

McCarthy, John J. 2000. Faithfulness and prosodic circumscription. *Optimality Theory: Phonology*, *Syntax*, *and Acquisition*, ed. by Joost Dekkers, Frank Reinoud Hugo van der Leeuw & Jeroen Maarten van de Weijer, 151-189. Oxford & New York: Oxford University Press.

McCarthy, John J., and Alan S. Prince. 1990. Foot and word in prosodic morphology: the Arabic broken plural. *Natural Language and Linguistic Theory* 8.2:209-283.

McCarthy, John J., and Alan S. Prince. 1993. Prosodic morphology I: constraint interaction and satisfaction. Manuscript. Amherst: University of Massachusetts; New Brunswick: Rutgers University.

McCarthy, John J., and Alan S. Prince. 1994a. Two lectures on prosodic morphology. Handouts of two lectures at OTS/HIL workshop on Prosodic Morphology, Utrecht University, July 1994. Available at Rutgers Optimality Archive (http://roa.rutgers.edu/, ROA#59).

McCarthy, John J., and Alan S. Prince. 1994b. The emergence of the unmarked. *Proceedings of the 24th Annual Meeting of the North East Linguistic Society* (*NELS 24*), ed. by Mercè Gonzàlez, 333-379. Amherst: GLSA.

McCarthy, John J., and Alan S. Prince. 1995. Faithfulness and reduplicative identity. *University of Massachusetts Occasional Papers in Linguistics 18: Papers in*

Optimality Theory, ed. by Jill N. Beckman, Suzanne Urbanczyk & Laura Walsh Dickey, 249-384. Amherst: GLSA.

Moravcsik, Edith. 1978. Reduplicative constructions. *Universals of Human Language*, Vol. 3: *Word Structure*, ed. by Joseph H. Greenberg, 297-334. Stanford: Stanford University Press.

Piñeros, Carlos E. 1998. *Prosodic Morphology in Spanish: Constraint Interaction in Word-Formation*. Columbus: Ohio State University dissertation.

Prince, Alan S. 1990. Quantitative consequences of rhythmic organization. *Chicago Linguistic Society* (*CLS*) 26.2:355-398. Chicago: Chicago Linguistic Society.

Prince, Alan S., and Paul Smolensky. 1993[2004]. *Optimality Theory: Constraint Interaction in Generative Grammar*. Malden: Blackwell.

Rose, Sharon, and Rachel Walker. 2004. A typology of consonant agreement as correspondence. *Language* 80.3:475-531.

Schein, Barry, and Donca Steriade. 1986. On geminates. *Linguistic Inquiry* 17.4:691-744.

Shen, Chia-chi. 2005. *Reflexives and Reciprocals in Kavalan*. Taipei: National Taiwan University MA thesis.

Smolensky, Paul. 1993. Harmony, markedness, and phonological activity. Handout of talk presented at Rutgers Optimality Workshop-1. New Brunswick: Rutgers University. Available at Rutgers Optimality Archive (http://roa.rutgers.edu/, ROA#87).

Spaelti, Philip. 1997. *Dimensions of Variation in Multi-Pattern Reduplication*. Santa Cruz: University of California dissertation.

Yip, Moira. 1995. Identity avoidance in phonology and morphology. Manuscript. Irvine: University of California.

Yu, Alan C. L. 2005. Quantity, stress and reduplication in Washo. *Phonology* 22.3:437-475.

Yu, Alan C. L. 2007. *A Natural History of Infixation*. Oxford & New York: Oxford University Press.

[Received 5 March 2012; revised 15 June 2012; accepted 6 August 2012]

Department of English
National Taiwan Normal University
162, Sec. 1, Hoping East Road
Taipei 106, Taiwan
hslin@ntnu.edu.tw

噶瑪蘭語重疊詞的變異

林蕙珊
國立臺灣師範大學

本論文以優選理論研究噶瑪蘭語重疊詞結構。噶瑪蘭語重疊詞綴有相當多樣的結構。前人研究噶瑪蘭語重疊詞時指出，該語言重疊詞綴的結構深受詞基首音節結構之影響。本文提出不同的看法。本文指出，噶瑪蘭語重疊詞受制於兩股衝突的力量；重疊詞一方面希望完整複製詞基首音節，另一方面又希望達到雙音拍 (bimoraic) 大小；兩股力量作用之下，使得有些詞基有一種以上的重疊結構。當詞基首音節為雙音拍時，重疊詞綴直接複製詞基首音節而無變異；而當詞基首音節小於雙音拍時，除了複製詞基首音節外，重疊詞綴也會複製詞基首音節以外的成分以達成雙音拍大小；此時，重疊詞綴便有一種以上的重疊結構。

關鍵詞：噶瑪蘭語，重疊詞，變異，優選理論