SOUTHERN VIETNAMESE REDUPLICATION: AN OPTIMALITY THEORY ANALYSIS

By

Nghi Lam

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Linguistics—Master of Arts

2025

ABSTRACT

This thesis offers a unified account of reduplication in Southern Vietnamese, concentrating on two under-documented polysyllabic patterns — *triplication* and *AABB* — and on the avoidance effects that block certain base–reduplicant combinations. Earlier studies describe these facts, but either restrict attention to disyllables or cannot capture how tone, avoidance, and reduplicant shape interact.

This study addresses these gaps through three main contributions. First, it provides substantial new data on previously under-documented reduplication patterns that (1) enhance Vietnamese linguistic documentation and (2) provide a robust empirical foundation for formal analysis. Second, it offers a description of the identical patterns avoided across reduplication types, previously sparsely covered in the literature. Third, I offer a formal analysis of the polysyllabic reduplication types — triplication and AABB — using tools from Optimality Theory (OT). Although the analysis draws on OT for formalization, it remains framework-neutral and broadly applicable.

A key claim of this thesis is that the so-called triplication in Vietnamese results from the attachment of *two distinct reduplicative morphemes*, each with its own allomorphs, and *not* recursive reduplication. In this view, Vietnamese reduplication is best understood as the use of *lexically specified* reduplicant morphemes. Specifically, morphology selects an allomorph according to phonological properties of the base (notably tone); phonology then merely enforces its placement. High-ranking alignment constraints and MAX-IO constraints keep the selected allomorph at its designated edge and preserve its lexical tone. Avoidance patterns therefore emerge from the *selectional restrictions* imposed by lexically specified allomorphs, thus avoiding theoretical issues such as ranking paradoxes.

Rather than invoking multiple constraint rankings, all attested patterns can be captured within a single unified grammar in which lexically specified reduplicants select appropriate stems. This allomorphy-based analysis provides a principled account of complex reduplication patterns in Southern Vietnamese, addressing and resolving previous theoretical challenges.

ACKNOWLEDGEMENTS

I am profoundly grateful to Professor Karthik Durvasula, my advisor, for his guidance, encouragement, and insightful feedback on every chapter of this thesis. Whenever self-doubt crept in, his reassurance kept me moving forward. I also thank my committee members, Professor Yen-Hwei Lin and Professor Silvina Bongiovanni, whose constructive comments and questions sharpened the analysis of Vietnamese triplication.

This work benefited from lively discussions with members of the Phonology Reading Group and my peers in the Department of Linguistics. I am also grateful to the linguistics faculty, whose courses have reminded me time and again how interesting linguistics is. Special thanks to Pam Withers, Ben Lampe, and Melissa French for patiently handling my many administrative questions and miscellaneous requests.

Last but not least, I owe deepest gratitude to my parents for their unwavering support through every decisions that I made. I thank my boyfriend, Dang, for his love and constant encouragement from miles away during the hardest stretches of this journey. I am equally grateful for my friends in Vietnam, especially Phuong, Lu, Thu, Van, and Lam, whose humor and support lift me up throughout the periods when everything feels impossible.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

Reduplication is a productive morphophonological method in Vietnamese which helps form a large amount of new words (Hoang et al., 2008). Consequently, it is one of the most extensively studied topics in Vietnamese linguistics (Hoang, 1985; Thompson and O'Harrow, 1988; Hoang, 1998; Cao, 2007; Vu, 2007; Hoang et al., 2008; *inter alia*). Most existing work consists of descriptive analyses of reduplication (Ngo, 1984; Hoang, 1985; Thompson & O'Harrow, 1988; D.-H. Nguyen, 1997; Srichampa, 2002). Other studies cover experiments and acoustic analyses of Vietnamese reduplication (T. Nguyen & Ingram, 2006; T. Nguyen & Ingram, 2007; A.-H. Pham & Pham, 2020). Formal analysis of Vietnamese reduplication, however, is limited (Agbayani, 1997; Inkelas & Zoll, 2005; Vu, 2007).

Among existing studies, disyllabic reduplication predominates in the literature, as it constitutes the majority of reduplicative words in Vietnamese. It embodies the key phonetic and semantic characteristics of Vietnamese reduplication, such as tone harmony and the encoding nuances of intensification and attenuation (Diep & Hoang, 1996). However, this has led to a research gap between disyllabic reduplication and other reduplication forms, such as *triplication* and *AABB*, even though they are also productive and have patterns that are similar to those of disyllabic reduplication.[1] Of these polysyllabic forms, triplication poses challenges in Vietnamese linguistics (P.-P. Nguyen, 1998; Hoang et al., 2008). However, little formal work has been done to analyze these challenges under a theoretical framework.

An important recurring pattern across Vietnamese reduplication involves *identity avoidance*, where speakers systematically avoid adjacent identical syllables. Despite its prevalence in Vietnamese and many world languages (Kennard, 2004; An, 2012; Wivell et al., 2024 among others), identity avoidance has not been adequately formalized in previous studies on Vietnamese reduplication, receiving only brief mention in Vu (2007).

This thesis makes three primary contributions. First, it provides substantial new empirical

[1]The notion of *productive* will be explained in Section 2.5.2.

data on previously under-documented reduplication patterns — triplication and AABB, used in tandem with the data and example sentences drawn from prior work by native speakers Ngo (1984), Hoang (1998), and Vu (2007). The novel data are obtained from my fieldwork, where informants provided acceptability judgments on various reduplicative forms and participated in an elicitation task in which informants were given a real word in the language with no known occurrence of triplication or AABB. This new dataset not only enhances the Vietnamese linguistic documentation but also provides the empirical foundation necessary for testing theoretical claims about reduplication processes in this thesis. Any instances of such novel data will be explicitly marked with *(Novel data)* throughout the thesis.

Second, this thesis proposes a formalization of Vietnamese reduplication patterns within Optimality Theory (OT), introducing a novel theoretical perspective: Vietnamese triplication is shown to result from the attachment of *two distinct reduplicative morphemes*, each with its own set of allomorphs, rather than repeated reduplication as previously suggested in the literature. Morphology chooses which allomorph surfaces, triggered by phonological properties of the base (notably tone), while phonology only enforces that choice through high-ranking Align and Max-IO that keep the allomorph at its stored edge and preserve its lexical tone.

Third, identity avoidance effects are extensively described, and are argued to result from ordinary *selectional restrictions*: an allomorph that would duplicate the base's rhyme is simply absent from the lexicon. No special anti-identity constraint, and no ranking paradox, is required. In other words, certain reduplicant allomorphs simply do not select stems of particular phonological shapes, and this lexical restriction superficially resembles identity avoidance. This approach removes the necessity of positing explicit phonological constraints specifically designed to avoid identity.

Existing analyses rely on *cophonologies* (Orgun, 1996; Inkelas et al., 1997; Anttila, 2002; Inkelas & Zoll, 2007) and *Morphological Doubling Theory*, both of which assume different

reduplication types have separate phonological grammars. In contrast, I propose a *unified* phonology where Vietnamese reduplication processes are best understood through *lexically specified reduplicant morphemes* and *allomorph selection*. Note that while the analysis is presented within the OT framework, it does *not* depend exclusively on it; rather, OT serves as a formalization tool for my proposal.

The rest of this thesis is organized into five chapters. Chapter 2 covers the relevant morphological and phonological aspects of Vietnamese linguistics. Descriptions of Vietnamese reduplication processes are given in Chapter 3. Chapter 4 addresses some existing approaches to analyses of triplication and AABB in Vietnamese and cross-linguistically. Chapter 5 goes through formal analyses provided for each reduplication process in Vietnamese. The thesis concludes with Chapter 6.

CHAPTER 2

BACKGROUND

This chapter discusses aspects of Vietnamese phonology and morphology relevant to the current analysis. Since the author and informants are Vietnamese speakers of the *Southern* variety, background information (except the syllable structure, word stress, and word structure) and phonetic transcriptions in the subsequent sections are of the Southern variety as well, unless mentioned otherwise. This is why the title explicitly refers to Southern Vietnamese reduplication.

This section is organized as follows: Section 2.1 covers the tone system of Southern Vietnamese, the final consonant and vowel inventories are briefly introduced in Section 2.2 and Section 2.3 respectively, Section 2.4 covers the syllable structure, Section 2.5 outlines different word structures in Vietnamese, and Section 2.6 closes out this section with a brief discussion on word stress in Vietnamese.

2.1 Tone

There are five tones in Southern Vietnamese.[1] The following description is from Brunelle (2009) accompanied by a graph in Figure 2.1. A1 is a level tone that starts around mid F0 range and has no creakiness nor breathiness throughout its production.[2] A2 is also a level tone, but it starts lower than A1 and is produced with a falling pitch. B1 is a rising tone but, similar to A1, has no creakiness nor breathiness throughout its production. B2 is a low falling-rising tone, and unlike its Northern counterpart, there is no final glottalization. C1 is a falling-rising tone with no laryngealization, similar to its Northern counterpart. Note that, unlike the Northern variety, there is no C2 in Southern Vietnamese since this is merged with C1.[3]

[1]There are different arguments for whether Vietnamese has six or eight tones, see A.-H. Pham (2001) and A. H. Pham (2004) for the discussion. For this analysis, the six-tone system is adopted.

[2]Alphanumerical labels for tones are adapted from Michaud (2005).

[3]See Brunelle (2009) for a detailed discussion of the production and perception of C1 and C2 in Northern and Southern Vietnamese.

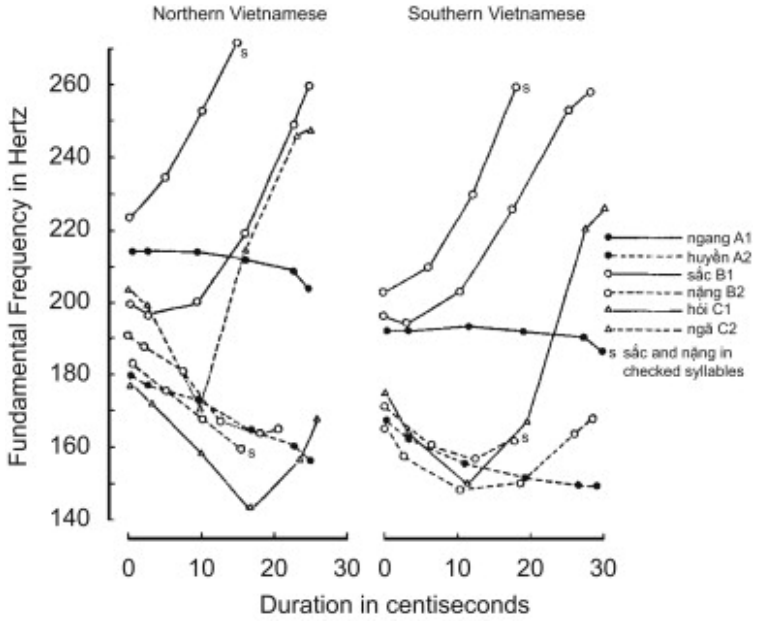


Figure 2.1 Average F0 contours of Northern and Southern Vietnamese tones (Brunelle, 2009)

A summary of the tone system is presented in Table 2.1. The organization is from Cao (2003); the contour description is adopted from A.-H. Pham (2001)'s formal structure of Vietnamese tones. Tones are organized in terms of their contours and registers. Register here refers to the vocal range where the tone is produced (Vo, 1997). There are two registers — high (+H) and low (−H). This means that tones A1, B1, and C1/C2 are produced at a higher vocal range. Tones A2 and B2 on the other hand are produced at a lower vocal range. Contour refers to the tonal shape; there are two general patterns — flat and non-flat (Vo, 1997; A.-H. Pham, 2001). Flat refers to level tones, with their overall shape not changing much throughout the tone. Non-flat, on the other hand, refers to tones that change their shape over the production time.

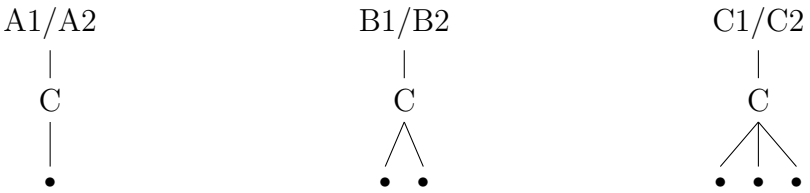
		Contour		
		Flat	Non-flat	
Register	+H	A1	B1	C1
	−H	A2	B2	

Table 2.1 Southern Vietnamese tones (A.-H. Pham, 2001; Cao, 2003)

Following A.-H. Pham (2001), the tone contour can be captured through the use of what the author refers to as *points*, denoted by a dot •. These points indicate whether there is movement during the tone's time course. A single point indicates no movement during the tone. If there is movement during the tone, more than one points are used to indicate the contour. The specification for tone contour of each tone is presented in (1).

The model in (1) has some of the following indications. During the time course of flat tone (i.e., A tone), there is no movement, hence one point under the C(ontour) node. There is movement in the non-flat tones (i.e., B and C tones); hence, there are two to three points under the C node. The difference between the number of points in the C node for either non-flat tone is irrelevant in the case of reduplication and thus will not be discussed in further detail.[4]

(1) Specification for tone contour



Based on their structural complexity and occurrence in Vietnamese, tone markedness is identified by register and contour (Vo, 1997; A.-H. Pham, 2001; A. H. Pham, 2004; Hoang et al., 2008). In terms of register, low tones are more marked than high tones. In terms of contour, the tone markedness is evaluated in pairs between flat and non-flat tones. C1-C2 are more marked than B1-B2 tones, which are more marked than A1-A2 tones. Thus, A1 is the most unmarked tone, while C1/C2 is the most marked tone.[5]. The markedness of tones is summarized in (2)–(3), where '<' indicates the left-hand tone is less marked than the right-hand one.

(2) Tone markedness in terms of register:

A1, B1, C2 < A2, B2, (C2)

(3) Tone markedness in terms of contour:

A1-A2 < B1-B2 < C1(-C2)

[4]See A.-H. Pham (2001) for a more in-depth analysis of Vietnamese tone.

[5]Under an eight-tone system, in terms of contour, B1-B2 in checked syllables (i.e., *D1* and *D2*) are the most marked; consequently, D2 is the most marked in Vietnamese (A.-H. Pham, 2001).

2.1.1 Tone-coda co-occurrence restriction

In East and Southeast Asian languages, there is a prevalent constraint on possible tone and coda combinations (Hyman, 2014). In the case of Vietnamese, this can be observed when the syllable has an *obstruent* coda. Specifically, B1 and B2 tones are allowed to appear in syllables with obstruent codas /p, t, k/ (e.g., (4a)) but A and C tones are not (e.g., (4b)).

(4) a. *thích* [tʰɨtB1] 'to like' b. * *hãc* [haːkC2]

2.2 Final consonants

In Vietnamese, only six consonants from the whole Vietnamese consonant inventory are allowed as coda. Note that /w/ is not usually counted as a coda consonant because it is considered to be an off-glide in diphthongs. The permissible codas are given in the following Table 2.2.

		Place			
		Labial	Alveolar	Palatal	Velar
Manner	Plosive	p	t		k
	Nasal	m	n		ŋ
	Approximant (Semivowels)	w		j	

Table 2.2 Vietnamese codas

Other permissible codas are the labial-velars [k͡p, ŋ͡m]. These are doubly articulated consonants that only occur after rounded vowels in Vietnamese, hence are allophones of /k, ŋ/ when following rounded vowels.

2.3 Vowels

There are eleven monophthongs in Vietnamese, which are presented in Table 2.3 below. Note that /i/ has an allophone [ɨ] that occurs with alveolar codas.

		Frontness		
		Front	Central	Back
Height	High	/i/		/ɯ/, /u/
	Mid	/e/	/ɤ/, /ɤː/	/o/
	Low	/ɛ/	/a/, /aː/	/ɔ/

Table 2.3 Vietnamese monophthongs (Thompson & O'Harrow, 1988)

Vowel (and consonant) features are given based on the feature geometric model proposed by Clements and Hume (1995) to account for the place specifications of consonants and vowels. The revised model of the V-place node from Parker (2018) is used instead of the Aperture node to distinguish the difference between central and back vowels. The revised model is given in Figure 2.2 below.

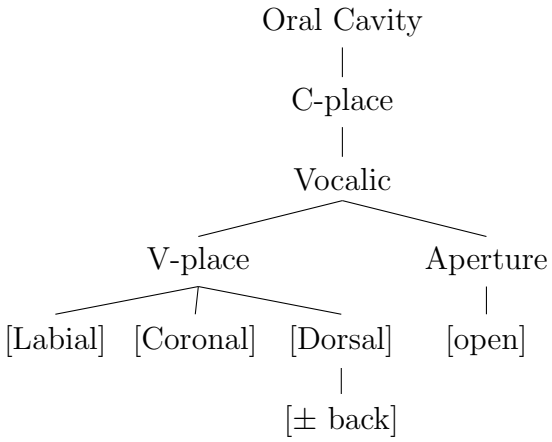


Figure 2.2 Feature geometric model (Clements & Hume, 1995)
and Revised model of the V-place (Parker, 2018)

Labial, coronal, and dorsal features are distinctive, in which they are not specified for a value {+, −, 0}. Front vowels /i, e, ɛ/ are defined as [coronal]. Central and back vowels are both specified as [dorsal]. To distinguish between central and back vowels, central vowels are defined as [−back], while back vowels are defined as [+back]. Among the back vowels, the rounded ones /u, o, ɔ/ are, in addition, specified with [labial]. For the analysis at hand, however, I will employ the binary feature nodes [±high] and [±low] instead of [±open$_{1,\ 2,\ 3}$] of the Aperture node to represent the vowel height. Thus, the high vowels will have the height nodes [+high, −low]; mid vowels will have [−high, −low], low vowels will have [−high, +low].

2.4 Syllable structure

The syllable structure in Vietnamese is given in Figure 2.3. In this structure, onset (C_1) and nucleus (V) are obligatory; on-glide (G_1), coda (C_2), and off-glide (G_2) are optional. Tone is realized on the rhyme (A.-H. Pham, 2001). There are arguments around the status of the on-glide, but I'll assume that the on-glide is a segment of its own and is a part of the rhyme (Cu et al., 1977; V.-L. Nguyen & Edmondson, 1997; A.-H. Pham, 2009).

Onset	Rhyme		
	Glide[6]	Nucleus	Coda
C_1	G_1	V	C_2/G_2[7]

Figure 2.3 Vietnamese syllable structure

2.5 Word structure

The organization of different word structures here mainly follows Ngo (1984), D.-H. Nguyen (1997), and Vu (2007). The smallest meaningful constituent possible is a *simple word*.[8] It is a word that contains only one morpheme that has a lexical or grammatical meaning, which can be divided further into *monosyllabic* (e.g., (5a)) and *polysyllabic* (e.g., (5b)) (D.-H. Nguyen, 1997).

(5) a. Monosyllabic word: [ɲaːA2] 'house(s)', [kɯəC1] 'door(s)', ...

b. Polysyllabic word: [ɓuA2.ɲɨnA2] 'scarecrow', [luŋ͡mA1.lɨnA2] 'sparkling', ...

Any words that are made up of two or more morphemes are *complex words*, which are divided into two groups — *compounds* (Section 2.5.1) and *reduplicatives* (Section 2.5.2).

2.5.1 Compounds

A compound is composed of two simple words and can be divided into two types — *coordinate* and *subordinate* (D.-H. Nguyen, 1997). Coordinate compound is composed of two

[6]Also called *on-glide* to differentiate from *off-glide* which occurs in coda, or *pre-vocalic* /w/ since only /w/ occurs here.

[7]Also called *off-glide*.

[8]Any labels used here for different word structures are from D.-H. Nguyen (1997).

words and has a *broader* meaning than that of its components. For instance, the word for *clothes* in Vietnamese (6a) is composed of two simple words: one for *shirts* and one for *pants*. Subordinate compound is composed of two words as well, but its meaning is *narrower* than that of its components, specifically, the head of the compound (the leftmost syllable). For instance, the word for *bicycle* in Vietnamese (6b) is composed of two simple words: one for *vehicle* and one for *pedal*; the composition of these two words results in a compound with a meaning that is narrower than that of the head of the compound, i.e., the word for *vehicle*.

(6) a. [ʔaːwB1] 'shirt(s)' + [kwɤŋA2] 'pants' = [ʔaːwB1.kwɤŋA2] 'clothes'

b. [sɛA1] 'vehicle' + [ɗaːpB2] 'pedal' = [sɛA1.ɗaːpB2] 'bicycle'

2.5.2 Reduplicatives

Reduplicatives in Vietnamese can be split into two groups: *productive reduplicatives* (*reduplicatives*, henceforth) and *lexicalized reduplicatives* (P.-P. Nguyen, 1998; Vu, 2007; A.-H. Pham & Pham, 2020). Productive reduplicatives are outputs of a reduplication process, and thus the main data of the analysis. On the other hand, lexicalized reduplicatives are *not* results of a reduplication process. They are listed in dictionaries, while productive ones are not because there is a large amount of them in the language and their constructions can be subsumed under a set of formation rules (P.-P. Nguyen, 1998).

A reduplicative has two components — a *base*, which carries the core meaning, and a *reduplicant*, which is the copy of the base that is affixed to it. To be specific, reduplicant here the actual phonological projection of a phonologically-unspecified *lexical entry* RED (McCarthy & Prince, 1994). An example is given below in (7). In this example, the reduplicant RED precedes the base. More reduplicative data will be presented in Chapter 3.

(7) [cɤmA2 cɤmB2]
RED slow
'Slightly slow'

Note that in the case of lexicalized reduplicatives, there are forms that can be broken down to obtain its base (and hence its structure is somewhat similar to that of a subordinate

compound) (e.g., (8a); the supposed base is boldface); however, there are other cases in which they can't be broken down any further (and hence similar to polysyllabic simple word) (e.g., (8b)). More examples of this reduplication form and explanations are offered in Section 3.6 and Appendix B.

(8) a. [luəŋA1.**luəŋA1**]
'always'

b. [ɗuA1.ɗuC1]
'papaya'

Additionally, reduplicatives are not counted as compounds because if we strictly follow the definition for compounds in this case, they are composed of simple words, while reduplicatives are not made up of simple words, but rather of morphemes (which happens to have simple words in the mix).

2.6 Word stress

In Vietnamese, there's little evidence for word stress. There is past literature that describe word stress in Vietnamese as final stress (i.e., word final); however, this can also be accounted as phrase-final lengthening and not word stress per se (Brunelle, 2017). However, experimental work on word stress in reduplication pointed out that the second syllable in disyllabic reduplication is more acoustically prominent (T. Nguyen & Ingram, 2006). There's also no acoustic evidence to support the claim of contrastive stress patterns between compounds and phrases. Thus, under forced elicitation condition, there's a juncture between constituents in phrases, but none in compounds (T. Nguyen & Ingram, 2007). As a consequence, each constituent in phrases bears stress, and constituents are produced continuously with the rightmost syllable stressed (T. Nguyen & Ingram, 2007; A.-H. Pham, 2009). For this analysis, I will assume that in Vietnamese, *the rightmost syllable* in the word will get the prominence stress, following T. Nguyen and Ingram (2007), Vu (2007), and A.-H. Pham (2009).

2.7 Interim summary

In this chapter, the descriptive groundwork was laid out for analyzing and understanding Vietnamese reduplication. The section started with examining the tonal system, coda consonants, and vowels, which are central to understanding the primary alternations between syllables in reduplicatives. Next, syllable structure was discussed to pinpoint where such alternations would occur. I discussed the word structure to clarify how the type of base influences and is influenced by the reduplication process. The chapter closes with word stress in Vietnamese to show its role in distinguishing the base from the reduplicant in a prosodic structure of the reduplicative. With these foundational elements in place, the next chapter will present and describe the reduplication data.

CHAPTER 3

DATA

This study focuses on productive reduplication processes. Thus, the following sections present data from reduplicatives derived from clearly replicable processes, including nonce forms. As mentioned in the introduction, all of the data, except for novel words from my informants, is from Ngo (1984), Hoang (1998), Vu (2007), and myself. Since the transcriptions were given for Standard (Northern) Vietnamese, I re-stated and re-described the data here with Southern transcriptions instead to align with the variety of Vietnamese used here.

Instead of categorizing reduplication forms by the number of syllables or the shape, I organized them by their morpho-semantic function: Section 3.1 for attenuation, Section 3.2 for exclamation, Section 3.3 for pejorative, Section 3.4 and Section 3.5 for intensification of different base and reduplication form, and a brief discussion of lexicalized reduplication forms in Section 3.6. To make the data presentation clear, the leftmost word in each example or data point is the base, the rightmost word is the reduplicative.

3.1 Attenuation

The predominant function of disyllabic reduplication is *attenuation*. This function is only applicable to adjectives and verbs.[1] In disyllabic reduplication, the reduplicant's form can either be a *full* or *partial* copy of its base. One thing to note is that every base can undergo both full *and* partial reduplication. However, the choice and frequency of each type depend on speaker preference.[2] In addition, attenuation reduplication limits to monosyllabic bases. If a speaker wants to attenuate the meaning of an adjective or verb that has more than one syllable (e.g., (1a) & (1b) respectively), the meaning has to be paraphrased.

(1) a. [luŋm͡A1.lɨnA2] ↛
sparkling
'sparkling'

b. [ŋiA1 ŋɤːA2] ↛
suspect suspect
'to suspect'

[1]Note, however, that some sources refer to the meaning of verbal (full) reduplication as verbs with iterative marking (Thompson & O'Harrow, 1988).

[2]See A.-H. Pham and Pham (2020) for SVN speakers' full vs. partial reduplication preference.

In the full reduplication case, the reduplicant is an exact copy of the base (e.g., (2a)).[3] In the partial reduplication case, the reduplicant copies all but the base's tone contour (and coda) (e.g., (2b) & (2c)).

(2) a. [kaːwA1] 'tall' — [kaːwA1 kaːwA1] / *[kaːwA2.kaːwA1] — RED tall — 'slightly tall'

b. [saːkB1] 'close to' — [saːŋA1 saːkB1] / *[saː**kA1**.saːkB1] — RED close.to — 'slightly close to'

c. [ɲɔC1] 'small' — [ɲɔA1 ɲɔC1] / [ɲɔC1 ɲɔC1] — RED small RED small — 'slightly small'

However, unlike adjectives, verbs can only take full reduplication. Consider the data given in (3). Partial reduplication in (3a) being unacceptable is expected because the base has a flat tone (i.e., A1), meaning regardless of the form, the reduplicant's tone is always a full copy of that of the base. The same partial form in (3b) is unacceptable as well, even though it obeys the tone alternation for partial reduplication and the tone-coda restriction.

(3) a. [ŋiA1] 'to suspect' — [ŋiA1 ŋiA1] / *[ŋiA2.ŋiA1] — RED suspect — 'to suspect(ATT)'[4]

b. [tʰɨtB1] 'like' — [tʰɨtB1 tʰɨtB1] / *[tʰɨnA1.tʰɨtB1] — RED like — 'to like(ATT)'

The degree of the verb is attenuated in the sense that the verb is less direct or blunt (e.g., (4a) with (3a) verb, where the base form appears in the sentence, compared to the corresponding (4b), where the reduplicative is used; base and reduplicant are in boldface). Similarly, attenuation also conveys uncertainty (e.g., (5a) with (3b) verb, compared to the (5b) reduplicative counterpart). However, not all verbs are compatible with attenuation, as

[3] In cases of full reduplication, the base and the reduplicant are identical in form, making it impossible to systematically distinguish between them. Thus, gloss annotations are assigned arbitrarily.

[4] ATT = attenuation

it is unusual to attenuate a dynamic verb like *eat* in (6), where (a) contains the base form, and (b) the reduplicative. Regardless, a discussion of the syntactic, semantic, and pragmatic properties of different verb types is beyond the scope of the current study.

(4) a. taːwA1 ɗaːC2 **ŋiA1** nɔB1 rojA2
1SG PAST suspect 3SG already
'I've suspected him/her/it for a while.'

b. taːwA1 ɗaːC2 **ŋiA1 ŋiA1** nɔB1 rojA2
1SG PAST RED suspect 3SG already
'I've kinda suspected him/her/it for a while.'

(5) a. taːwA1 **tʰɨtB1** nɔB1
1SG like 3SG
'I like him/her/it.'

b. taːwA1 **tʰɨtB1 tʰɨtB1** nɔB1
1SG RED like 3SG
'I kinda like him/her/it.'

(6) a. ʔaːŋA1 **ʔaŋA1** kɤmA1
An eat rice
'An has a meal.'

b. # ʔaːŋA1 **ʔaŋA1 ʔaŋA1** kɤmA1
An RED eat rice
'An kinda has a meal.'

3.1.1 Tone pattern

The tone harmony pattern in partial reduplication is *same-register* tone harmony, in which the reduplicant's tone contour is always A with its register copied from the base. One explanation for why it is an A tone is because of *tone neutralization*; the reduplicant might have been a total copy of the base, but the tone is neutralized to an *unmarked* tone (i.e., a flat (A) tone) (Hoang et al., 2008). Due to the tone-coda co-occurrence restriction, if the base has an *obstruent* coda (e.g., (2b)), the reduplicant will not take the base's coda but will have the homorganic nasal instead.

Another possible explanation, that is of interest of the analysis in Chapter 5, is *fixed flat tone* in the lexical representation of the reduplicant. This means that the reduplication morpheme has a fixed flat tone, with the function of attenuation in the language.

A consequence of the fixed A tone contour and full copy of the base's tone register is that if the base has a flat tone (e.g., the unacceptable form in (2a)), the reduplicant is always an exact copy of the base (cf. the acceptable and the unacceptable form in (2a)). These two

tone harmony patterns are considered to be the representative tone pattern in Vietnamese reduplication as a whole (Diep & Hoang, 1996; Cao, 2007).

3.2 Exclamation (AABB)

In past literature, AABB (in this analysis, exclamation) reduplication has been described to be used for repetition (for adjectives and verbs) (Srichampa, 2002) and emphatic (Vu, 2007). (7) shows examples of how the base form and its reduplication for exclamation are in sentences.[5] Compared to (7a), sentence (7b) contains the reduplicative. This form emphasizes the large amount of clothes piling up. Consequently, the sentence as a whole functions as an exclamation..

Unlike attenuation and pejorative reduplication, exclamation can only take disyllabic inputs. If we have a monosyllabic base (e.g., [ʔaːwB1] 'shirt(s)') like (7c), we cannot have a sentence like (7d) that has the reduplication of [ʔaːwB1] 'shirt(s)'. Instead, we can re-phrase the sentence like (7e), which is the same as (7c) with an additional question particle (Q) [vɤjB2], changing the sentence to a rhetorical question that expresses exclamation.

(7) a. ʈɔŋ͡mA1 fɔŋ͡mA2 A **ʔaːwB1 kwɤŋA2** cɤkB1 t^{h}anA2 nujB1
in room A shirt pants stack into mountain
'In A's room, clothes pile up like a mountain.'

b. ʈɔŋ͡mA1 fɔŋ͡mA2 A **ʔaːwB1 ʔaːwB1 kwɤŋA2 kwɤŋA2** cɤkB1 t^{h}anA2 nujB1
in room A shirt shirt pants pants stack into mountain
'In A's room, clothes pile up like a mountain. (EMPH)'[6]

c. ʈɔŋ͡mA1 fɔŋ͡mA2 A **ʔaːwB1** cɤkB1 t^{h}anA2 nujB1
in room A shirt stack into mountain
'In A's room, shirts pile up like a mountain.'

d. * ʈɔŋ͡mA1 fɔŋ͡mA2 A **ʔaːwB1 ʔaːwB1** cɤkB1 t^{h}anA2 nujB1
in room A shirt shirt stack into mountain
'In A's room, clothes pile up like a mountain. (EMPH)'

e. ʈɔŋ͡mA1 fɔŋ͡mA2 A **ʔaːwB1** cɤkB1 t^{h}anA2 nujB1 vɤjB2
in room A shirt stack into mountain Q
'Why are there so many shirts in A's room?'

[5] (7) is novel data.
[6] EMPH = emphasis

Similar to pejorative reduplication, this form applies to items across all lexical categories. For adjectives, there is an alternation between this reduplication and the [-aː] reduplication to express emphasis or intensification of the adjective's degree.

A common feature among all lexical categories is that the input must take the form of a *coordinate compound* (e.g., (8a), (9a), and (10b)). For adjectives, lexicalized reduplicatives are also possible inputs (e.g., (10a)). Additional novel data (see Appendix B) further confirms that lexicalized reduplicatives can serve as valid inputs across categories (e.g., (8b), (9b), and (10c)).

By contrast, inputs with other structures, such as subordinate compounds, reduplicatives, and mono- or polysyllabic words, are not possible. This includes nominal and adjectival subordinate compounds (e.g., (8d), (10d)), nominal and verbal lexicalized reduplicatives (e.g., (8d), (9c)), nominal polysyllabic words (e.g., (8b)), verb phrases (e.g., (9d)), and adjectival productive reduplicatives (e.g., (10d)).

(8) **Nouns**:

a.	[ʔaːwB1 kwɤŋA2] shirt pant 'clothes'		[ʔaːwB1 ʔaːwB1 kwɤŋA2 kwɤŋA2] shirt shirt pant pant 'clothes (EMPH)'	
b.	[ʔaːwB1 ʔiəkB1] shirt $R_{iək}$ 'shirts (PEJ)'		[ʔaːwB1 ʔaːwB1 ʔiəkB1 ʔiəkB1] shirt shirt $R_{iək}$ $R_{iək}$ 'shirts (PEJ, EMPH)'	(Novel data)
c.	[kaːA2.fɛA1] 'coffee'	*	[kaːA2.kaːA2.fɛA1.fɛA1] *Intended*: 'coffee (PEJ)'	
d.	[sɛA1.ɗaːpB2] 'bicycle'	*	[sɛA1.sɛA1.ɗaːpB2.ɗaːpB2] *Intended*: 'bicycle (PEJ)'	
e.	[cuəŋA2.cuəŋA2] 'dragonfly'	*	[cuəŋA2.cuəŋA2.cuəŋA2.cuəŋA2] *Intended*: 'dragonfly (PEJ)'	

(9) **Verbs**:

a.	[kɯəjA2 nɔjB1] laugh talk 'to laugh and talk'	[kɯəjA2 kɯəjA2 nɔjB1 nɔjB1] laugh laugh talk talk 'to laugh and talk (EMPH)'	
b.	[ʔaŋA1 ʔiəkB1] eat $R_{iək}$ 'eat (PEJ)'	[ʔaŋA1 ʔaŋA1 ʔiəkB1 ʔiəkB1] eat eat $R_{iək}$ $R_{iək}$ 'eat (PEJ, EMPH)'	(Novel data)
c.	[ŋuC1.ŋeA1] 'sleep'	* [ŋuC1.ŋuC1.ŋeA1.ŋeA1] *Intended*: 'sleep (PEJ, EMPH)'	
d.	[laːmA2.ɓaːjA2] 'to study'	* [laːmA2.laːmA2.ɓaːjA2.ɓaːjA2] *Intended*: 'to study (PEJ, EMPH)'	

(10) **Adjectives**:

a.	[mɤA1.maːŋA2] dreamy 'dreamy'	[mɤA1 mɤA1 maːŋA2 maːŋA2] dream dream - -[7] 'dreamy (EMPH)'	
b.	[sɨnA1 ɗɛpB2] pretty pretty 'pretty'	[sɨnA1 sɨnA1 ɗɛpB2 ɗɛpB2] pretty pretty pretty pretty 'pretty (EMPH)'	
c.	[sopB1 siəkB1] fluffy $R_{iək}$ 'fluffy (PEJ)'	[sopB1 sopB1 siəkB1 siəkB1] fluffy fluffy $R_{iək}$ $R_{iək}$ 'fluffy (PEJ, EMPH)'	(Novel data)
d.	[vaːŋA2.vaːŋA2] 'slightly yellow'	* [vaːŋA2.vaːŋA2.vaːŋA2.vaːŋA2] *Intended*: 'slightly yellow (EMPH)'	
e.	[sanA1.laːB1] 'green'	* [sanA1.sanA1.laːB1.laːB1] *Intended*: 'green (EMPH)'	

[7] [maːŋA2] doesn't have meaning in Vietnamese, but it's an [-aːŋA] reduplicant of an Adj, an unproductive reduplication process

Similar to pejorative reduplication, this reduplication form also exhibits separability effect, shown in (11).[8] The phrase [vəjB1.caːC1] 'also NEG' can be used as the "separator". Same conclusion can be made from the separability effect for exclamation form, that is the base and RED in exclamation reduplication belongs to separate prosodic structures, i.e., each is a prosodic word in the same phonological phrase.

(11) **ʔaːwB1 ʔaːwB1** vəjB1 caːC1 **kwɤŋA2 kwɤŋA2** muəA1 jiA2 maːA2 ɲiəwA2 vɤjB2
shirt shirt with NEG pants pants buy what then many Q
'You bought so many clothes, and for what? (PEJ, EMPH)'

The general form of the reduplicative is AABB, where each syllable of the input AB is reduplicated. Because the copies of A and B are segmentally identical to A and B themselves, it is difficult to determine which instance is the RED. There is no empirical evidence at this point to definitively identify which element is the reduplicant in AABB forms, since both copies are segmentally identical. I have chosen to analyze the leftmost syllable of each pair as the reduplicant, making them *prefixes*, primarily for consistency with other reduplication processes in Vietnamese, particularly the attenuation forms, which are generally analyzed as prefixal. This assumption allows for a more uniform treatment across different reduplication types in the language.

Exclamation reduplication bears a resemblance to the *AABB reduplication* found in Chengdu (Lin, 2015) and Taiwanese (Hsiao, 2018), particularly regarding its form and the way it intensifies meaning. For this reason, I will extend their analyses of alignment in AABB reduplication to the Vietnamese data.

3.3 Pejorative ([-iək])

Pejorative reduplication is highly productive; this form can be applied to any lexical item, regardless of structures or number of syllables, in Vietnamese. Unlike attenuation, in pejorative reduplication, the reduplicant always follows the base. For example, in (12a), the

[8] (11) is novel data.

reduplicant [siəkB1] is preceded by its base [sɛA1] 'vehicle' and not the other way around (i.e., *[siəkB1.sɛA1]).

(12) a. [sɛA1] vehicle 'vehicle(s)' — [sɛA1 siəkB1] / *[sɛA1.siəkB2] / *[siəkB1.sɛA1]; vehicle $R_{iək}$; 'vehicles (PEJ)'[9]

b. [vɤːC1] notebook 'notebook(s)' — [vɤːC1 viəkB1] / *[vɤːC1.viəkB2]; notebook $R_{iək}$; 'notebooks (PEJ)'

c. [ɓaːŋB2] friend 'friend(s)' — [ɓaːŋB2 ɓiəkB2] / *[ɓiəkB1/B2.ɓaːŋB2]; [ɓaːŋB2 ɓiəkB1]; friend $R_{iək}$; 'friends (PEJ)'

The reduplicant's rhyme is always [-iək]; it never copies the base's rhyme. If the base has two syllables or more (e.g., (13a)–(g)), the number of syllables in $R_{iək}$ has to be the same as that of the base.[10] The rightmost syllable then will be the one that has the [-iək] rhyme; any syllables preceding $R_{iək}$ in the reduplicative will copy their correspondents in the base. For instance, in (13a), given the base [kaːA2.fɛA1] 'coffee', $R_{iək}$ copies both syllables but changes the rhyme of the rightmost syllable to [-iək]. Hence, we have the reduplicant [kaːA2.fiəkB1].

(13) a. [kaːA2.fɛA1] coffee 'coffee(s)' — [kaːA2.fɛA1 kaːA2.fiəkB1 / *B2]; coffee $R_{iək}$; 'coffees (PEJ)'

b. * kiəkB2 fɛA1
(number of σ's in the reduplicative ≠ number of σ's in the base)

c. * kaːA2 fiəkB1 kaːA2 fɛA1
($R_{iək}$ in the second syllable)

d. * kiəkB2 fɛA1 kaːA2 fɛA1
($R_{iək}$ in the first syllable)

e. * kaːA2 fɛA1 kiəkB2 fɛA1
($R_{iək}$ in the third syllable)[11]

[9] PEJ = pejorative
[10] $R_{iək}$ = reduplicant in pejorative process
[11] These violations are the same inputs of any lexical categories.

f.	[ʔaːwB1 kwɤŋA2] shirt pant 'clothes'	[ʔaːwB1 kwɤŋA2 ʔaːwB1 kiəkB1/B2] shirt pant shirt $R_{iək}$ 'clothes (PEJ)'	
g.	[kɤwA1.lakB2.ɓoB2] club 'club(s)'	[kɤwA1.lakB2.ɓoB2 club kɤwA1.lakB2.ɓiəkB2/B1] $R_{iək}$ 'clubs (PEJ)'	(Novel data)

According to Vu (2007), pejorative reduplication is only possible for bases with at most two syllables. This is false, as shown by cases like (13g) in which the base has three syllables and the process is still applicable, as long as characteristics of pejorative reduplication (fixed rhyme, tone contour, etc.) are obeyed.

One characteristic of pejorative reduplication that will be crucial for the analysis of the prosodic structure of the reduplicative in the complete work is the *syntactic separability* effect (Vu, 2007). This means that additional syntactic constituents can separate the base from $R_{iək}$ (e.g., (14)–(15), where the base and reduplicant are in boldface; for more examples, see Appendix B).[12] Specifically, in (14), (a) contains the base, while (b)–(c) contain the reduplicative. In (15), (a)–(b) contain the base, while (c)–(f) contain the reduplicative. This effect also occurs in coordinate compounds and lexicalized reduplicatives. The fact that the base and $R_{iək}$ can be separated suggests that each belongs to a separate prosodic structure, i.e., each is a prosodic word within the same phonological phrase (Vu, 2007).

(14) **Monosyllabic base**: Separating constituent (SEP) = [xoŋm͡A1] 'no'

a. ɛmA1 xoŋm͡A1 ciwB2 **ŋuC1** jiA2 hɤtB1
2SG NEG agree sleep what all
'You never sleep or anything at all!'

b. ɛmA1 xoŋm͡A1 **ŋuC1 ŋiəkB1** jiA2 hɤtB1
2SG NEG sleep $R_{iək}$ what all
'You never sleep or anything at all! (PEJ)'

[12]Syntactic structures of the constituents used as the "separator" are left open for future research as they are beyond the scope of this analysis.

c. X SEP $\overline{\text{X}}$

ɛmA1 xoŋ͡mA1 **ŋuC1** xoŋ͡mA1 **ŋiəkB1** jiA2 hɤtB1
2SG NEG sleep NEG $\text{R}_{\text{iək}}$ what all
'You never sleep or anything at all! (PEJ)'

(15) **Polysyllabic base (coordinate compound)**: SEP = [vəjB1.caːC1] 'also NEG'

a. **kɯəjA2 nɔjB1** (kaːjB1) jiA2 ɗɔB1
laugh talk (CLF) what that
'What are you laughing and talking about?'

b. X SEP Y

kɯəjA2 vəjB1 caːC1 **nɔjB1** (kaːjB1) jiA2 ɗɔB1
laugh with NEG talk (CLF) what that
'What are you laughing and talking about? (PEJ)'

c. **kɯəjA2 nɔjB1 kɯəjA2 niəkB1** (kaːjB1) jiA2 ɗɔB1
laugh talk laugh $\text{R}_{\text{iək}}$ (CLF) what that
'What are you laughing and talking about? (PEJ)'

d. X Y SEP $\overline{\text{X}}$ $\overline{\text{Y}}$

kɯəjA2 nɔjB1 vəjB1 caːC1 **kɯəjA2 niəkB1** (kaːjB1) jiA2 ɗɔB1
laugh talk with NEG laugh $\text{R}_{\text{iək}}$ (CLF) what that
'What are you laughing and talking about? (PEJ)'

e. * X SEP Y $\overline{\text{X}}$ $\overline{\text{Y}}$

kɯəjA2 vəjB1 caːC1 **nɔjB1 kɯəjA2 niəkB1** (kaːjB1) jiA2 ɗɔB1
laugh with NEG talk laugh $\text{R}_{\text{iək}}$ (CLF) what that
Intended: 'What are you laughing and talking about? (PEJ)'

f. * X Y $\overline{\text{X}}$ SEP $\overline{\text{Y}}$

kɯəjA2 nɔjB1 kɯəjA2 vəjB1 caːC1 **niəkB1** (kaːjB1) jiA2 ɗɔB1
laugh talk laugh with NEG $\text{R}_{\text{iək}}$ (CLF) what that
Intended: 'What are you laughing and talking about? (PEJ)'

3.3.1 Tone pattern

Since the reduplicant has a fixed [-iək] rhyme and the tone-coda restriction (see Section 2.1), there is only one possible tone for the reduplicant, that is B tones. As for the tone register, it is said to be a same-register tone harmony for Southern (and Central) speakers, unlike the Northern variety in which $\text{R}_{\text{iək}}$'s tone register is always +H (i.e., the tone is always

B1). However, additional novel data (see Appendix B) shows that, for SVN speakers, there's an alternation between the use of B1 and B2 in $R_{iək}$, i.e., both B1 and B2 are possible in the reduplicant with any bases, and $R_{iək}$ does not have to copy the tone register of the base, regardless of the number of syllables in the base (e.g., (12c) and (13f)). There is a preference for B1 however, which would make it similar to the Northern variety's.

However, there is no tone alternation when the tone register of the base is +H; $R_{iək}$'s tone register has to be +H too (e.g., (12a)–(b), and (13a)). The description given in Vu (2007) is at least true for SVN when the base has two or more syllables; in such cases, B1–B2 alternation still exists, but there is a preference for both the base and $R_{iək}$ to have the same tone register. One possible explanation is that since +H tones are unmarked, there is no need for further alternation; any changes will result in a more marked tone. One problem with this is that there are no attested processes in Vietnamese that would favor high tones over non-high tones, and among reduplication processes themselves, there is no such a pattern. For this study, I will provide an analysis for both B1 and B2 in monosyllabic bases, and tone register copying in polysyllabic bases; the preference between B1 and B2 is a probable future research topic as it is not within the scope of the current study.

3.3.2 Avoidance pattern

When the base's rhyme is [-iək], the process is blocked. In that case, the pejorative meaning of the base can be expressed periphrastically or a replacement allomorph can be used instead. For instance, in (16b)–(a), a replacement allomorph, which carries the pejorative marking of the base, is used to express the pejorative meaning of the base.

(16) a. [siəkB1] — [siəkB1.suːŋm͡C1] / * [siəkB1 siəkB1]
circus — circus-PEJ — circus $R_{iək}$
'circus(es)' — 'circuses (PEJ)'

b. [tiəkB2] — [tiəkB2.tuːŋm͡A2] / * [tiəkB2 tiəkB2] / *[tiəkB2.tiəkB1]
party — party-PEJ — party $R_{iək}$
'party(s)' — 'parties (PEJ)'

In cases where the base has the rhyme [-iək] with tone B1, as in (16a), speakers reject a form like *[siəkB1.siəkB1]. One plausible reason is *homophony avoidance* in which such a string is identical to the obsolete productive *full* nominal reduplication pattern that means 'every circus'. To avoid that, speakers resort to a periphrastic form shown in (16a).

The same story cannot explain [-iəkB2] bases (16b). Even if the reduplicant switched to B1, *[tiəkB2.tiəkB1], the string is *not* homophonous with any existing word, yet it is still rejected. What matters is the *phonological (partial) identity* of the rhyme, in which the patterns avoided are not completely identical to each other (only the rhyme [-iək] (and the tone contour but the tone register). Vietnamese therefore instantiates a more general ban on adjacent *identical rhymes*, rather than simple homophony avoidance.

Partial-identity avoidance of this sort is cross-linguistically common. French noun-forming *-iste* /ist/ blocks /ist/ after a base that already ends in [-is] or [-iz] (e.g., /bodis/ "Baudis (name)" + /-ist/ → [bodist], *[bod**isis**t]; see Corbin and Plénat (1992) and de Lacy (1999)). Nisgha shows a similar restriction: the third-person singular suffix *-t* /t/ cannot follow stems ending in [s], [t], or [ɬ] (Russell, 1995; de Lacy, 1999). Moreover, the morphological identity effects surveyed in Yip (1995) — English plural *-s* + possessive *'s* and Chinese Mandarin double *-le* (perfective marker *-le* + sentence-final particle conveying current state change *-le*) — also block adjacent identical material, though those languages repair the violation by deletion of either copy of the affix. Another way to consider this is that there is a *null allomorph* in the lexicon that emerges when there are adjacent identical segments (i.e., English) or affixes (i.e., Chinese Mandarin).

By contrast, no repair is necessary in Vietnamese because there is simply *no* null allomorph in the lexicon, unlike cases in Chinese Mandarin and English.. The process is ineffable when there are adjacent identical segments (i.e., [-iək] base + [-iək] reduplicant). Morphology therefore selects a periphrastic alternative as the reduplicant, and phonology merely enforces that choice. The avoidance effect is thus *morphological*, a matter of allomorph inventory, not a phonological identity constraint. Treating it this way avoids the ranking paradoxes that

an explicit anti-identity constraint would create (see Section 5.5).

3.4 Intensification (1)

In Vietnamese, there are two reduplication processes that can be used to express intensification of a polysyllabic verb or an adjective — one form involves a fixed rhyme of [-aː] and the other form is triplication.[13] This section discusses the first process.

[-aː] reduplication is only applicable to *disyllabic* verbal and adjectival reduplicatives. Existing data shows that this form is observed with lexicalized reduplicatives (e.g., (17)); however, novel data shows that reduplicatives from productive processes, pejorative to be specific, are also possible (e.g., (18a)–(c)). As a result, this process is possible for nouns as well (e.g., (18a)).

(17) a. [kaŋA2.ɲaŋA2]
'complain'
[kaŋA2.ɲaːA2.kaŋA2.ɲaŋA2]
'constantly complain'

b. [lɨnA1.tɨnA1]
'nonsensical'
[lɨnA1.taːA1.lɨnA1.tɨnA1]
'nonsensical (EMPH)'

c. [sɤːŋB1.saːkB1]
'ill-mannered'
[sɤːŋB1.saːA1.sɤːŋB1.saːkB1]
'ill-mannered (EMPH)'

(18) a. [ɲaːA2 ɲiəkB1]
house $R_{iək}$
'houses (PEJ)'
[ɲaːA2 ɲaːA1 ɲaːA2 ɲiəkB1]
house $R_{aː}$ house $R_{iək}$
'houses (PEJ, EMPH)'
(Novel data)

b. [ɣɛpB1 ɣiəkB1]
couple $R_{iək}$
'to couple (PEJ)'
[ɣɛpB1 ɣaːA1 ɣɛpB1 ɣiəkB1]
couple $R_{aː}$ couple $R_{iək}$
'to couple (PEJ, EMPH)'
(Novel data)

c. [vaːŋA2 viəkB1]
yellow $R_{iək}$
'yellow (PEJ)'
[vaːŋA2 vaːA1 vaːŋA2 viəkB1]
yellow $R_{aː}$ yellow $R_{iək}$
'yellow (PEJ, EMPH)'
(Novel data)

[13] To better distinguish the two intensification processes, I will refer to the former process as **[-aː] reduplication** and the latter as **triplication**, utilizing their main segmental difference.

Similar to the pejorative process, this reduplication also involves a fixed rhyme and tone contour, [-aː] and A tone contour, respectively. The tone harmony pattern is also of the same register, similar to reduplication processes described above.

The general form is $\overline{\text{AB}}$AB in which $\overline{\text{B}}$ is $\text{R}_{\text{aː}}$ with fixed segments, and $\overline{\text{A}}$ is RED of A, $\overline{\text{B}}$ for B.[14] Thus, this form resembles attenuation and exclamation reduplication, in that the reduplicant is a prefix.

This reduplication also exhibits syntactic separability effect, presented in (19).[15] Here, the reduplicative [kaŋA2.ɲaːA2.kaŋA2.ɲaŋA2] 'constantly complaining' is split by the word [hajA1] 'or'.

(19) xoŋm͡ **kaŋA2.ɲaːA2** hajA1 **kaŋA2.ɲaŋA2** jiA2 hɤtB1
NEG RED or complain what all
'Stop complaining whatsoever!'

3.4.1 Avoidance pattern

When the base's rhyme is [-aː] and has an A tone, the process is blocked. Having either the rhyme [-aː] or the flat tone alone does not block the reduplication, as shown by (20) with the rhyme [-aː]. But when the two patterns simultaneously occur in the base, the process is blocked, as shown in (21) (Vu, 2007). In such cases, the intensification meaning can be expressed periphrastically or the speaker can use a replacement form instead. For instance, given a base like [lɤwA1.laːA1] '(of time) long' in' (21), the base (the rightmost syllable, to be specific) has [-aː] as its rhyme and an A tone, so this reduplication process cannot be applied; a replacement form is used instead. In this case, the replacement form is obtained through the exclamation reduplication process (Vu, 2007).

(20) [ɓɤjB2.ɓaːB2] [ɓɤjB2.ɓaːA2.ɓɤjB2.ɓaːB2] (Novel data)
'inappropriate' 'very inappropriate'

[14] Since the reduplicant's rhyme is always [-aː], for convenience, I will refer to the reduplicant in intensification (1) as $\text{R}_{\text{aː}}$.

[15] (19) is novel data.

(21) [lɤwA1.laːA1] [lɤwA1 lɤwA1 laːA1 laːA1]
long RED long RED -[16]
'(of time) long' '(of time) long (EMPH)'

* [lɤwA1 laːA1 lɤwA1.laːA1]
long $R_{aː}$ long
'*Intended*: '(of time) long (EMPH)'

The same allomorph-selection logic from Section 3.3 and Section 5.5 applies here, which will be discussed in Section 5.6. This also parallels with English and Chinese Mandarin identity avoidance (Yip, 1995).

3.5 Intensification (2)

Triplication is a productive process in Vietnamese, yet it is documented in a limited amount of literature (Hoang et al., 2008). It has been described and generalized in only a few studies (see Appendix B; cf. Hoang, 1985, 1998; P.-P. Nguyen, 1998; Srichampa, 2002; Hoang et al., 2008). Moreover, triplication, with a similar intensifying function, can also be found in various languages or dialects, such as Taiwanese (Zhang & Lai, 2007) and the Chengdu variety of Mandarin (Lin, 2015).

Based on these available data (a sample list of data given in (22)), triplication in Vietnamese applies to only *monosyllabic* adjectives. In the reduplicative, the first syllable is assumed to be the base because the segments and the tones of the monosyllabic input are preserved here (Hoang et al., 2008). This syllable is also said to carry the primary stress of the word (Hoang et al., 2008); (P.-P. Nguyen, 1998) extended the stress carrier to include the third syllable as well. The second syllable consistently bears an A2 tone and, if the base ends in an *obstruent* coda, it also has a homorganic *nasal* coda, due to the tone-coda co-occurrence constraint.

For the third syllable, there are two sub-patterns: if the base has an B1 tone and is a checked syllable (i.e., one that has an *obstruent* coda), then the third syllable has B2 (cf. (22a)–(b) and (22c)); otherwise, the reduplicant will have an A1 tone (and homorganic nasal

[16]Since the syllable does not have a meaning, a dash - is used in the gloss.

coda if there is a coda in the base). In general, the third syllable and the first syllable (i.e., the supposed base) harmonize in terms of either contour (i.e., flat versus non-flat tones) tones or register (i.e., high versus low tones) (Hoang et al., 2008).

(22)

a.	[saːkB1] 'close to'	[saːkB1.saːŋA2.saːkB2] 'very close to'	b. [ʔɨtB1] 'little'	[ʔɨtB1.ʔɨnA2.ʔɨtB2] 'very little'[17]
c.	[sɤwB1] 'ugly'	[sɤwB1.sɤwA2.sɤwA1] 'very ugly'	d. [satB2] 'clean'	[satB2.sanA2.sanA1] 'very clean'
e.	[tɯŋC1] 'crazy'	[tɯŋC1.tɯŋA2.tɯŋA1] 'very crazy'	f. [lwaːŋC2] 'watery'	[lwaːŋC2.lwaːŋA2.lwaːŋA1] 'very watery'

Besides the triplication pattern presented in (22), another pattern of form X$\overline{\text{Y}}$Y that has *three syllables* is also documented, shown in (23). In these cases, the input is an adjectival phrase of the form XY; the adverb (i.e., the rightmost syllable Y) is the base for reduplication. The tone harmony pattern here is similar to that in attenuation. However, they function differently — while the same pattern in attenuation is for degree minimizing, it is for intensification here.

(23)

a.	[xoA1.xok͡pB1] 'very dry'	[xoA1.xoŋ͡mA1.xok͡pB1] 'very very dry'
b.	[tʰɤːmA1.fɯkB1] 'very aromatic'	[tʰɤːmA1.fɯŋA1.fɯkB1] 'very very aromatic'
c.	[satB2.ɓatB1] 'very clean'	[satB2.ɓanA1.ɓatB1] 'very very clean'

In the current analysis, the focus on triplication is of the type in (22), in which there are *two* copies of the one base, and not the type in (23) in which there is *only one* copy of the base. The tone pattern of X$\overline{\text{Y}}$Y can then be subsumed under that of partial reduplication presented in Section 3.1.

[17] Novel data

A number of triplication forms have a truncated disyllabic form, in which the second syllable in the triplication form is omitted from the disyllabic one (e.g. (24); for more examples, see Hoang et al. (2008)). For instance, (22a)/(24a) can be truncated to [saːkB1.saːkB2] 'very close to' (i.e., (24e)); both forms yield the same meaning, but the triplication form has slightly more emphasis. However, this truncation is not possible in all cases of triplication, as shown by the illicit forms in (24g)–(h).

(24) a. [saːkB1.saːŋA2.saːkB2]
'very close to (EMPH)'

b. [jɯŋC1.jɯŋA2.jɯŋA1]
'indifferent (EMPH)'

c. [hɔkB1.hɔŋA2.hɔkB2]
'very hot'

d. [tɯŋC1.tɯŋA2.tɯŋA1]
'very crazy'

e. [saːkB1.saːkB2]
'very close to'

f. [jɯŋC1.jɯŋA1]
'indifferent'

g. * [hɔkB1.hɔkB2]
'very hot'

h. * [tɯŋC1.tɯŋA1]
'very crazy'

3.5.1 Tone pattern

In general, there are five tonal schema (base-reduplicant alignment asides) observed in triplication: (1) B1.A2.A1, (2) B1.A2.B2, (3) B2.A2.A1, (4) C1.A2.A1, and (5) C2.A2.A1, in which each tone counts for a syllable in the reduplicative. Of these patterns, (2), (3), and (5) do not follow tone harmony within the same register if we consider the first syllable and the third syllable to be the base and the reduplicant respectively.

Consider (25a) for schema (2) and (25b) for schema (3). In (25a), the base and the second reduplicant only differ in the tone, where both tones B1 and B2 are non-flat. However, they belong to *different registers* (i.e., B1 being +H and B2 being −H) and thus do not follow the same-register harmony pattern. If both of the reduplicants and the base follow the tone pattern in partial reduplication (see Section 3.1), the first reduplicant (i.e., the second syllable) should have been [saːŋA1] and [saːŋA1] for the second one (i.e., the third syllable) as well. However, these two forms will (1) result in identical syllables in a sequence, and (2)

changing the fixed A2 in the second syllable, which is illicit in Vietnamese triplication (see Section 5.8.1 for the discussion).

In (25b), assuming the first syllable [satB2] to be the base, its relations with both the first reduplicant [sanA2] and the second reduplicant [sanA1] are different from each other. The second reduplicant does not follow the harmony pattern; it is not in the same register nor has the same contour as that of the base (i.e., A1 in the reduplicant and B2 in the base). If this reduplicant and the base were to follow the tone pattern in partial reduplication, the second reduplicant should also be [sanA2], i.e., having an A2 tone, which is the flat counterpart of B2, but that would result identical syllables in a sequence.

(25) a. [saːkB1 saːŋA2 saːkB2]
close RED RED
'very close to'

b. [satB2 sanA2 sanA1]
clean RED RED
'very clean'

In conclusion, both tonal harmony patterns in examples of triplication in (25) are different from each other and different from the general pattern of tonal harmony in partial reduplication. This pattern also leads to another problem in triplication — base and reduplicant alignment, described in the following section.

3.5.2 Avoidance pattern

Triplication is blocked when the base tone is A2. In such cases, the intensification meaning has to be conveyed by paraphrase or through replacement forms. For instance, in (26), an existing word in the language that means 'very naughty' is used in place of the triplicated form of [liA2] 'naughty'.

(26) [liA2] 'naughty' [liA2.lɤmB2] 'very naughty' / *[liA2.liA2.liA1/B2] (Novel data)

If A2 bases were to be allowed, given the existing tone pattern in triplication, there would be two possible schema: (a) A2.A2.A1 and (b) A2.A2.B2. In the schema (b), which has B2 in the third syllable, the base may or may not have an obstruent coda, but it has A2 tone and not B1; thus, this schema is impossible. The schema (a) is not possible as well. Section 5.8

will go into the details of why a form like schema (a) is not possible, in which this is also due to *allomorphy selection*, similar to pejorative and intensification (1).

3.5.3 Reduplicant placement

When the base's tone is A1, the base is no longer the first syllable, but is now the third syllable instead, shown in the boldface syllable in (27). The difference in the placement of the base leads to the shift of the morphemes R_{A2} and RED. The reduplicants follow the base for all cases, in which R_{A2} immediately follows the base and then RED. The order is in reverse for A1 base; R_{A2} precedes the base and them the RED precedes R_{A2}. Additionally, the tone in the reduplicant is also changed. In (27), the reduplicant has an C1 tone instead of A1. Keeping the base in the first syllable and changing the tone of the reduplicant in the third syllable from A1 to C1 is not possible too.

(27) [nɤwA1] 'brown' — [nɤwC1 nɤwA2 **nɤwA1**] (Novel data)
RED R_{A2} brown
'very brown'

* [nɤwA1 nɤwA2 **nɤwA1**] RED R_{A2} brown
Intended: 'very brown'

* [**nɤwA1** nɤwA2 nɤwC1] brown R_{A2} RED
Intended: 'very brown'

Note that there is a difference between the base in (28a)–(b) and that in (27). The analysis for this mismatch in alignment will be presented Section 5.8, in which I will argue for these two reduplicants being *two separate morphemes*, each having different allomorphs that are subject to their own alignment constraints.

(28) a. [**jɯŋC1**.jɯŋA2.**jɯŋA1**]
'very indifferent'

b. [**kɔŋC1**.kɔŋA2.**kɔŋA1**]
'very small'

3.6 Lexicalized reduplicatives

Lexicalized reduplicatives here are forms that are *not* produced by a productive word formation process in Vietnamese (P.-P. Nguyen, 1998; Vu, 2007). Example data are shown in (29)–(34). By the nature of the studies, most past literature listed and analyzed these

forms along with the productive reduplicatives and did not make a distinction between the two (Hoang, 1985; D.-H. Nguyen, 1997; Hoang et al., 2008). The "base" (if there's any) is determined by whether either component bears a lexical category in Vietnamese (e.g., (29)–(34)), but there are other forms that neither components of the reduplicatives bear any meaning (e.g., (35)).

Some forms, however, may appear like a result of a productive reduplication process. For instance, in the forms in (31), both constituents of the reduplicative are similar all but the onset, which can be considered as an example of partial reduplication. However, the onset alternation cannot be explained phonological nor phonetically. Additionally, this process is not observed elsewhere in the language, and it cannot be applied to other words as a word formation like the productive reduplication processes described in previous sections. Or words in (35a)–(b), they resemble outputs of partial reduplication (i.e., same-register tone harmony). However, they are *not* productive reduplicatives, and neither of the components can be the base. One explanation might be that these words historically derived from a productive process but that process nor the form itself is productive (Vu, 2007).

(29) (Full) reduplication of nouns (distributive and/or plurality marking Thompson and O'Harrow, 1988; Vu, 2007)

a. [aːjA1 aːjA1]
who who
'everyone'

b. [ŋɯjA2 ŋɯjA2]
person person
'everybody'

(30) (Full) reduplication of adverbs

a. [luəŋA1 luəŋA1]
always always
'always'

b. [hɤːjA1 hɤːjA1]
slightly slightly
'slightly'

(31) Onset replacement

a. [lanA1 canA1]
clever -[18]
'hasty'

b. [luəŋB1 kuəŋB1]
- panic
'in a hurry'

[18]Since the syllable does not have a meaning, a dash - is used in the gloss.

(32) Onset and tone replacement

a. [cɤːjA1 ɓɤːjA2]
play -
'to play around'

b. [satB2 ɓatB1]
clean -
'completely clean'

(33) Rhyme replacement

a. [kamB1 kujB1]
plant -
'to concentrate'

b. [ɓɯkB2 ɓojB2]
upset -
'upset'

(34) Rhyme and tone replacement

a. [maːkB1 mɛC1]
cool -
'(of weather) cool'

b. [tɨnC1 tawB1]
awake -
'refreshed'

(35) Cases in which the base and the reduplicant are indistinguishable

a. [tʰɨnA1.tʰɨtB1]
'silently'

b. [ɓiA2.ɓɔmC2]
'water sound'

3.7 Interim summary

In this chapter, a description for each of the productive reduplication processes is presented. In terms of shape, there are two main patterns: full and partial. Observations reveal that some reduplication processes but the exclamation one block forms with certain rhyme-tone combinations (e.g., [-iək] is blocked in the pejorative process or [-aːAα] is blocked in the exclamation process). In contrast, attenuation and exclamation, allow identical syllables. These are discovered through novel data and comparative examples across processes, showing systematic gaps. Moreover, since the scope of this thesis specifically targets the Southern variety rather than the Northern or Standard variety, I have clarified and re-described previous accounts of data where necessary. Preliminary analyses of the data have also been proposed and will be explored in detail in Chapter 5. In the following chapter, previous theoretical analyses of Vietnamese reduplication and similar processes in other languages will be reviewed.

CHAPTER 4

ANALYSIS REVIEW

Description and formal analyses of Vietnamese reduplication are provided in a number of works — Ngo (1984), Hoang (1985), Thompson and O'Harrow (1988), Agbayani (1997), D.-H. Nguyen (1997), Srichampa (2002), and Vu (2007). In this study, three representative work will be reviewed — descriptive (P.-P. Nguyen, 1998; Hoang et al., 2008) and formal analyses (Vu, 2007).

In addition, because the analysis of triplication and exclamation reduplication will draw on similar processes in other languages, a review of triplication in Chengdu Mandarin (Lin, 2015) and Taiwanese (Zhang & Lai, 2007), and AABB in Chengdu Mandarin (Lin, 2015) and Taiwanese (Hsiao, 2018) will be briefly discussed as well.

4.1 P.-P. Nguyen (1998)

P.-P. Nguyen (1998) discusses the morphological aspect of Vietnamese through two processes, *reduplication* and *affixation*. Description of reduplication is given from two approaches: the reduplicative's meaning and the reduplicating pattern.

In terms of meaning, there are two types. One is reduplication as a *grammatical* process in which the derived meaning is predicable and can be grouped into two types, *attenuation* and *intensification*. The other is a *lexical* process in which the derived meaning is *not* predictable from the reduplicating pattern. In general, this is similar to the classification shown in Chapter 3.

Pattern-wise, there are two types: *onset alternation* and *tone alternation*. Onset alternation refers to cases where the onsets in both syllables are changed; either or both the rhyme and the tone are preserved. For example, (1) shows [ʔ-] changing to [ɲ-] in the onset and (2) shows [ʔ-] to [t-] and [s-] (examples from Hoang et al. (2008)).

(1) a. [ʔeB1] 'lonely' [ʔeB1.ɲeB2] 'completely lonely'

b. [ʔiwB1] 'weak' [ʔiwB1.ɲiwB2] 'very weak'

(2) a. [ʔoŋm͡A1.toŋm͡A2]
'soft spoken'

b. [ʔɯmA1.sɯmA2]
noisy

Note that there has been no satisfactory explanation for the alternation between any of the onset pairs, and no fixed meaning is associated with the changes. Thus, this is classified as the lexical type because there is no predictable alternation pattern and the meaning associated to the pattern.

Tone alternation refers to cases where the tones of the syllables in the word can have either the same contour and register or just the register. There are two patterns — *full* and *partial*, shown in (3a)–(b) respectively. The derived meaning can be split into two groups, attenuation and intensification, shown in (3b)–(c) respectively. Note that in each of the examples, the base of the reduplication process is identifiable.

(3) a. [kaːwA1] [kaːwA1 kaːwA1]
tall RED tall
'tall' 'slightly tall'

b. [saːkB1] [saːŋA1 saːkB1]
close.to RED close.to
'close to' 'slightly close to'

c. [saːkB1] [saːkB1.saːŋA2.saːkB2]
'close to' 'very close to'

However, there are also cases where the meaning is lexicalized. For example, consider [ɗuA1.ɗuC1] 'papaya' and [ɓɨmA2.ɓɨpB2] 'coucal'. These two words have the tone harmony pattern of partial reduplication, but the base is *not* identifiable, unlike cases in (3a)–(c). In addition, neither of them have an attenuating or intensifying meaning. Thus, they are of the lexical type of reduplication.

Affixation covers the [-ɤp] affix with examples shown below. This is a prefix that signifies movement or instability of the verb it attaches to.

(4) a. [foŋm͡A2] [fɤpB1.foŋm͡A2]
'inflate' 'inflate and deflate'

b. [ŋɯŋA2] [ŋɤpB2.ŋɯŋA2]
'stop' 'hesitant'

It is similar to the pejorative process with the [-iək] rhyme discussed here in the sense that both have a fixed rhyme, takes the onset of the stem that it attaches to, and presents some tone harmony pattern. However, it is not a productive pattern like [-iək].

P.-P. Nguyen (1998) only gives a description of the tone harmony pattern in reduplication and affixation and how they are used to form new words in Vietnamese. The author does not give a formal analysis; however, they are one of the limited number of works that acknowledges triplication.

4.2 Hoang et al. (2008)

Hoang et al. (2008) discusses several morphological aspects of Vietnamese, including reduplication. Similar to P.-P. Nguyen (1998), they describe it as both a word formation process and a word type, but provide greater detail and cover more patterns. Reduplication is classified based on syllable count, derived meaning, and harmony pattern.

Of particular relevance here are the two hypotheses proposed by Hoang et al. (2008) for triplication: (1) repetition of a disyllabic base and (2) repetition of a monosyllabic base. The first treats triplication as reduplication of a disyllabic form, motivated by examples like [saːkB1.saːŋA2.saːkB2] 'very close to (EMPH)' in (5a), where the second reduplicant (A2) can be omitted, resulting in a disyllabic form with less emphasis (5b).

(5) a. [saːkB1 saːŋA2 saːkB2]
close.to RED RED
'very close to (EMPH)'

b. [saːkB1 ~~saːŋA2~~ saːkB2]
close.to RED RED
'very close to'

This leads to the question of whether triplication is derived by reduplicating a disyllabic base. If so, the process can be schematized as follows:

(6) $X.\overline{X} \rightarrow X.\overline{X}_{A2}.\overline{X}$

However, this hypothesis faces issues, which are many trisyllabic forms lack disyllabic counterparts, and the base–reduplicant alignment varies. Sometimes the base precedes the reduplicant (as in 5b), while in other cases it follows (28a)–(b). Additionally, no explanation is offered for why the reduplicant has A2. Due to these inconsistencies, Hoang et al. (2008) ultimately favors the second hypothesis.

In the alternative view, triplication begins with a monosyllabic base that can function as a word of its own. There are two structural possibilities: either the base and reduplicants

are contiguous (e.g., (7)) or not (e.g., (8)).

(7) a. $X \rightarrow X.\overline{X}_{A2}$

b. $X.\overline{X}_{A2} \rightarrow X.\overline{X}_{A2}.\overline{X}$

(8) a. $X \rightarrow X.\overline{X}$

b. $X.\overline{X} \rightarrow X.\overline{X}_{A2}.\overline{X}$

Hoang et al. (2008) posits that tone assignment follows arbitrary rules, which are the second syllable has fixed A2 tone, and the third syllable harmonizes with the base either by the tone contour or the tone register. Yet, this account mirrors the disyllabic-base hypothesis. Both suffer from inconsistent base placement and tone harmony. If this pattern can be proposed for the base and the third syllable in the trisyllabic form, the same thing can be proposed for the disyllabic-base hypothesis as well.

A notable difference between Hoang et al. (2008) and earlier works like P.-P. Nguyen (1998) and Vu (2007) (and this thesis) lies in lexicalized reduplication. The latter treat reduplication with onset/rhyme alternation as lexicalized; Hoang et al. (2008), however, sees them as compounds only if both syllables are meaningful. If only one or neither is meaningful, they are reduplicatives, even if non-productive, as long as there is a "harmony" between the two syllables in the word. An example is the [ʔ-]–[ɲ-] onset pair in (1); there is no reason for why this occurs or if this would make the syllables more harmonious (P.-P. Nguyen, 1998).

While both perspectives have merit, this thesis restricts its focus to productive reduplication patterns — those described in P.-P. Nguyen (1998) and Vu (2007) — to analyze patterns that are predictable. The scope also extends to triplication and AABB patterns, for which formal analyses will be proposed.

4.3 Vu (2007)

Vu (2007) provides a unified OT analysis for four productive reduplication processes: full, partial, [-iək], and [-aː] reduplication. The different processes are differentiated by their tone/segmental pattern: full, *Emergence of the Enmarked* (TETU) (i.e., partial) (McCarthy & Prince, 1995), *Melodic Overwriting* (i.e., [-iək] and [-aː] reduplication) (Alderete et al., 1999). Partial reduplication is TETU because it is the emergence of the unmarked tone —

flat tones. [-iək] and [-aː] reduplication are melodic overwriting because neither rhymes can be phonologically derived.

The analysis involves conflicting constraint rankings within a language; to account for that, the author opts for *Cophonology* (Orgun, 1996; Arto Anttila, 1997; Inkelas and Zoll, 2007 among others). This allows distinct co-existing phonological systems, or *cophonologies*, in a language. There's a *Master Ranking* which is the partial ranking of constraints that all individual cophonologies must obey.

The author proposes two cophonologies that the four reduplication processes follow: (1) *Marked ≫ Ident-BR and (2) Ident-BR ≫ *Marked in which *Marked is the set markedness constraints for the V(owel)-place. The two cophonologies share other constraints that are dominant in all reduplication processes: Faith-IO, Phono_{A} for tone markedness, alignment constraints, and FtBin. Note that alignment constraints refer to a set of alignment constraints that govern the placement of the base and the reduplicant in reduplicative; there are two different alignment constraints for the reduplicants, one [-iək] and one for all the remaining reduplicants, because only [-iək] is a suffixal reduplicant but others are prefixes. Three reduplication processes but the partial one belong to the Ident-BR ≫ *Marked cophonology. *Marked ≫ Ident-BR is the TETU ranking schema, and thus the partial reduplication belongs to this one for the unmarked tone (i.e., flat tones) to emerge.

Vu (2007) is one of the few that formalized the reduplication processes in Vietnamese. The analysis accounts for the different tone patterns in the processes as well as the prosodic structure of the base and the reduplicatives using cophonologies. However, Vu (2007) does not discuss AABB and triplication, even though they are also productive reduplication processes. In addition, there is no analysis given for the ineffability of [-iək] and [-aːAα] bases when they have the same rhyme as the fixed rhyme in the reduplicant of the pejorative and intensification processes. The author also provides that reduplicatives of the intensification process do *not* have the syntactic separability effect; however, novel data from my field work show that there is such an effect. That poses a possible problem for the analysis in Vu (2007)

such that one might have to propose an additional ALIGN constraint, i.e., ALIGN-L($R_{a:}$, ϕ), which rules that for every reduplicant $R_{a:}$, there is a phonological phrase such that the left edge of $R_{a:}$ coincides with the left edge of ϕ. This thesis will address and fill in the gaps listed here.

4.4 Chengdu Mandarin AABB and tone sandhi (Lin, 2015)

Lin (2015) examines four reduplication patterns in Chengdu Mandarin, *AA*, *AAB*, *ABB*, and *AABB*, and proposes an OT analysis to account for the reduplicant size and placement, and the tone sandhi of these patterns. For the purpose of this analysis, the focus will be solely on AABB, as exemplified below.

(9) a. si45.uən31 — si45.si$\overline{\mathbf{55}}$[1].uən31.uən31
'gentle' — 'very gentle'

b. ta13.faŋ**44**$^{<45}$ — ta13.ta$\overline{\mathbf{55}}$.faŋ**44**$^{<45}$.faŋ**44**$^{<45}$
'generous'[2] — 'very generous'

There are four possible ways to derive the AABB form from a disyllabic base AB, as shown below.

1. $A_1\overline{A_2B_1}B_2$
2. $A_1\overline{A_2}B_1\overline{B_2}$
3. $\overline{A_1}A_2B_1\overline{B_2}$
4. $\overline{A_1}A_2\overline{B_1}B_2$

Forms (3) and (4) are ruled out because the tone in the initial position is always identical to that in the input, suggesting that the initial syllable must be the base. The forms in (1) and (2) differ by how much of the reduplicant or the base intervenes one another. In (1), the base is intervened by the reduplicant, while in (2), not only is the base intervened by part of the reduplicant (A_2), the reduplicant is also intervened by part of the base (B_1). Thus, in terms of CONTIGUITY (McCarthy & Prince, 1995), (1) violates only CONTIG-IB in (10), while (2) violates both CONTIG-IB *and* CONTIG-IR in (11). Thus, the desired form is (1), $A_1\overline{A_2B_1}B_2$, in which the reduplicant appears to be disyllabic like its base and infixal.

[1] $\overline{\mathbf{T}}$ = special sandhi tone
[2] $\mathbf{T}^{<\mathrm{T}}$ = **sandhi tone** < citation tone

(10) CONTIGUITY-IB (CONTIG-IB): The portion of the base standing in correspondence to the input forms a contiguous string. (Fitzgerald, 2000)

(11) CONTIGUITY-IR (CONTIG-IR): The portion of the reduplicant standing in correspondence to the input forms a contiguous string. (Fitzgerald, 2000)

According to Lin (2015), reduplicants in different patterns are considered as allomorphs of the same functional affix, and the alternation in the reduplicant size and placement can be accounted for by *Morphoprosodic alignment* (MPA) model by Kennedy (2008). For the reduplicant size, it is proposed that the different sizes are the result of two separate patterns of reduplicants — *stem-internal* and *stem-external* — shown below.

(12) Stem-internal reduplicant: $[\text{RED} + \text{B}]_{\text{stem}}$

(13) Stem-external reduplicant: $[\text{RED} \# [\text{B}]_{\text{stem}}]$

Each of the patterns is accounted for by two alignment constraints, ALIGN + SYLL and ALIGN # FOOT, respectively. Based on MPA model, the disyllabic reduplicant in AABB is stem-external; thus, only ALIGN # FOOT is shown in (14).

(14) ALIGN # FOOT : Stem-external morphological boundaries occur at foot boundaries.

Although the reduplicant in AABB appears infixal on the surface, it is actually *suffixal* in nature.[3] Based on (14), there are two possible reduplicant alignment: (a) A.B. # $\overline{\text{A}}.\overline{\text{B}}$ and (b) A. # $\overline{\text{A}}.\overline{\text{B}}$ # .B. The desired form is (b), which is the same to the form proposed for AABB. It is triggered by a set of linearity constraints, LIN(M) in (15) and LIN(S) in (16), that requires the input segmental sequence to be preserved in the output.

(15) LIN(M): The linear order of the morphemes in the input is kept in the output.

(16) LIN(S): The linear order of the segments in the input is kept in the output.

The proposed ranking is in (17) with the input *ta13.faŋ**44***$^{<45}$ 'generous'. The tone is omitted from the tableau for simplicity. Candidate (c) is eliminated because it is not a foot

[3]This is similar to the other reduplication patterns in Chengdu Mandarin.

and hence violates the alignment constraint.[4] Candidate (a) of the form A.B. # $\overline{\text{A}}.\overline{\text{B}}$ is eliminated because it violates the linear order of segments, in which *ta* precedes *faŋ* in the input but it is not always the case in this output candidate (e.g., $\overline{ta}$ follows *faŋ*). This leaves candidate (b) of the form A. # $\overline{\text{A}}.\overline{\text{B}}$ # .B to be the optimal output at the expense of violating the linear order of the base and the reduplicant. This output does not violate Lin(S) because the segment *ta* preceding *faŋ* in the input is preserved here.

(17) Lin(S), Align # Foot ≫ Lin(M)

ta.faŋ # Red		Lin(S)	Align # Foot	Lin(M)
a.	ta.faŋ. # ($\overline{\text{ta.faŋ}}$)	∗!		
b. ☞	ta. # ($\overline{\text{ta.faŋ}}$) # .faŋ			∗
c.	ta.faŋ. # $\overline{\text{.faŋ}}$	∗!	∗!	

Having discussed the alignment of the base and the reduplicant, we now turn to the tone sandhi. The general tone sandhi rules are listed in Table 4.1.

	Input	Output	Example
a.	45.45	45.**44**	tɕin tʰiɛn ‘today’
b.	31.45	31.**44**	tso tʰiɛn ‘yesterday’
c.	13.45	13.**44**	miɛn pau ‘bread’
d.	55.45	55.**44**	tsʰau ku ‘straw mushroom’
e.	45.55	45.**53**	tsʰau səu ‘wonton’
f.	31.55	31.**53**	tʰɑŋ ko ‘candy’
g.	13.55	13.**53**	fu mu ‘parents’
h.	55.55	55.**53**	tso səu ‘left hand’
i.	45.13	45.**11**	tsoŋ tɕiau ‘religion’
j.	31.13	31.**11**	ɕi kuan ‘habit’
k.	13.13	13.**11**	sɨtɕiɛi ‘world’
l.	55.13	55.**11**	tsau fan ‘breakfast’

Key: tones are separated by ‘.’; tones that undergo general tone sandhi are in boldface.

Table 4.1 Tonal combinations that undergo tone sandhi in Chengdu (Lin, 2015)

[4]There is a dominant FtBin constraint which requires the foot to be minimally disyllabic.

These combinations can be summarized as rules in (18). Specifically, the tone 45 changes to tone 44 when preceded by another tone; the tone 55 shifts to tone 53 when in the phrase-final position; and the tone 13 changes to 11 when preceded by another tone. Notice that environment that triggers the change in tone 45 and 13 are similar to each other and hence the shaded rows in Table 4.1.

(18) General tone sandhi (TS) rules:

1. 45 [Hr, lh] → **44** [Hr, l]/T__[5]
2. 55 [Hr, h] → **53** [Hr, hl]/__]
3. 13 [Lr, lh] → **11** [Lr, l]/T__

In $A_1\overline{A_2B_1}B_2$ reduplication,[6] A_2 always bears the fixed tone $\overline{\mathbf{55}}$, while $\overline{B_1}$ and B_2 undergo regular tone sandhi as shown in Table 4.1 and encoded in (18). For instance, the word for 'tidy', tsən55.tɕʰi31, has the form tsən55.tsən$\overline{\mathbf{55}}$.tɕʰi31.tɕʰi31 'very tidy' when reduplicated.

(19) tsən55.tɕʰi31 'tidy' → tsən55.tsən$\overline{\mathbf{55}}$.tɕʰi31.tɕʰi31 'very tidy'

This fixed $\overline{\mathbf{55}}$ tone on the reduplicant is assumed to be a floating tone introduced by the morphological reduplication process. When the reduplicant is disyllabic, this floating H tone docks onto its left edge, a behavior governed by the constraints *Float, MaxFloat, and Align-HR-L, as defined below.

(20) *Float: No floating tones. (Wolf, 2005)

(21) MaxFloat: All autosegments that are floating in the input have output correspondents. (Wolf, 2005)

(22) Align(H, L, Red, L) (abbr. Align-HR-L): A high tone must be associated with the leftmost tone bearer of a reduplicant. (Carleton & Myers, 1994)

The desired ranking is in (23). Notice that since the lack of the realization of an input floating tone in the output would violate either *Float or MaxFloat, as in candidate (a).

[5] **T** = sandhi tone

[6] The subscript numbers are used for ease of reference.

Candidate (b) is eliminated because the input floating tone is realized, but it is associated with the rightmost tone bearer of a reduplicant instead.

(23) (T) = floating tone in the input; **[T]** = output tones that correspond to an input floating tone

45-RED, (55)	*FLOAT	MAXFLOAT	ALIGN-HR-L
a. 45-$\overline{\textbf{33}}^{7}$	*!	*!	
b. **[55]**-$\overline{\textbf{33}}$			*!
c. ☞ 45-$\overline{\textbf{[55]}}$			

Based on the general tone sandhi rules in (18), the allotone pairings 45∼**44**, 55∼**53**, and 13∼**11** all preserve tone register, with faithful preservation of the left tone segment from input to output. This supports the dominance of the constraints IDENT-IO-REG and IDENT-IO-L-T, shown in (24) and (25), respectively.

(24) IDENT-IO-REG— The register of an input tone must be preserved in the output.

(25) IDENT-IO-L-T — The left tone segments of an input must be preserved in the output.

The alternations from 45 and 13 to 44 and 11 in post-tonal or non-initial positions can be explained as leveling of rising contours and *deletion* of the left tone segment. This is captured by the constraints MAX-IO-T-L (no deletion at left edge) and *RISE (ban on rising contours), whose ranking is given in (26). This ensures that the change from 45, 13 to **44**, **11** occur in non-initial position only and *not* when 45 and 13 occur in isolation.

(26) MAX-IO-T-L ≫ *RISE

The shift from 55 to **53** in utterance-final position is attributed to a boundary low tone that incorporates with the base 55 tone at the end of the utterance. This is supported by the

[7]It is unclear why $\overline{\textbf{33}}$ is here. It might have been a typo in the original text because according to the general tone sandhi in Table 4.1, it should have been $\overline{\textbf{44}}$.

constraint Boundary Low, which enforces the parsing of boundary L tones at the surface (Chen, 2000). The ranking in (27) ensures that this alternation results from boundary tone incorporation rather than insertion of new tone segments.

(27) Boundary Low ≫ Dep-IO-T

Further evidence for this comes from how isolated tones 45 and 13 interact with boundary L tones. In such cases, the boundary L can surface either as a complex tone (i.e., 454[Hr, lh̲] and 131[Lr, lh̲]), in which the input tone is preserved, or a level tone (i.e., 44 [Hr, l̲] and 11[Lr, l̲]), in which the right input tone segment is deleted for the boundary L to surface.[8] The constraint *ComCon (ban on complex contours) is needed to prevent the former scenario, and its dominance over Boundary Low is shown in (28). Additionally, the change of 45 and 13 to **44** and **11** in the post-tonal/non-initial position, indicating that *Rise, which triggers the alternation, dominates Max-IO-T, which forbids the change.

(28) *ComCon, Max-IO-T ≫ Boundary Low

When combining the general and special sandhi processes in $A_1\overline{A_2B_1}B_2$ sequences, two cases arise: (1) when the base has 55 (i.e., A_1.55) and (2) when the base has 45 or 13 (i.e., A_1.45 or A_1.13). When the base has 55, it is copied onto the reduplicant $\overline{B_1}$. the B_1 in the input, which is now B_2 in the output, has 55 in the utterance's final position and thus is changed to **53** (i.e., $A_1.\overline{A_2}.\overline{55}.\mathbf{53}$) due to the boundary L. The change, however, is not reflected in the reduplicant, violating Ident-BR but obeying Boundary Low. If this change were copied, it would instead violate Ident-IR, and hence, the ranking proposed in (29). When the base has 45 or 13, the change to a level tone 44 or 11 is also applied to the reduplicant $\overline{B_1}$, which means the constraint *Rise that drives this change dominates Ident-IR.

[8] l̲ = boundary L tone

(29) IDENT-IR, BOUNDARY LOW ≫ IDENT-BR

X.55 # RED	IDENT-IR	BOUNDARY LOW	IDENT-BR
a. X-$\overline{\text{X.}\mathbf{53}}$-**53**	∗!		
b. X-$\overline{\text{X.55}}$-55		∗!	
c. ☞ X-$\overline{\text{X.55}}$-**53**			∗

Finally, a special tone sandhi condition arises when the reduplicant follows a base tone 55. In this context, the reduplicant never surfaces with any of the expected general sandhi tones (55, 13, 45, **11**, **44**, **53**), but instead with $\overline{\mathbf{31}}$. Tone 55 is avoided because the language militates against 55.$\boxed{55}$, in which $\boxed{55}$ corresponds to the floating H tone, violating (30). Tones 13 and 45, both having a rising contour LH, are universally disfavored due to *RISE. Tones 11 and 44 are ruled out by ¬BR-CON (31), since they share contours with 55 (i.e., 11: L, 44: L, and 55: H). Lastly, tone 53 is excluded because it shares a register with 55, which is High, violating ¬BR-REG (32).[9]

(30) *55.$\boxed{55}$: 55.$\boxed{55}$ is prohibited.

(31) ¬BR-CON: The reduplicant and the base should have different contours.

(32) ¬BR-REG: The reduplicant and the base should have different registers.

All of this comes to the ranking in (33). However, in the case of fixed 55 in $\overline{\text{A}_2}$ in $\text{A}_1\overline{\text{A}_2\text{B}_1}\text{B}_2$, the tone of $\overline{\text{B}_1}$ does not change to $\overline{\mathbf{31}}$ This means that IDENT-BR is preserved and hence the ranking in (34).

(33) *RISE, *55-$\boxed{55}$ ≫ ¬BR-CON, ¬BR-REG

(34) IDENT-BR ≫ ¬BR-CON, ¬BR-REG[10]

[9]Note that the anti-faithfulness constraints (31)–(32) are controversial due to limited cross-linguistic support and their data-driven formulation. A full evaluation lies beyond the scope of this summary.

[10]*RISE, *55-$\boxed{55}$ dominate IDENT-BR because *RISE ≫ IDENT-IR and *55-$\boxed{55}$ ≫ *FLOAT, MAXFLOAT ≫ IDENT-IR in other processes.

Lin (2015) offers *separate* analyses for reduplicant alignment and size on one hand, and tone sandhi on the other. As a result, there is no interaction between the two shown; neither constraint set appears to influence the rankings or outcomes of the other. While CONTIG-IB and CONTIG-IR are invoked to rule out unattested forms, their rankings relative to the tonal or alignment constraints are not explored in their paper. Additionally, the constraints in (31)–(32) are not easy to motivate beyond the data at hand. It is important to provide an analysis that is consistent between the reduplication patterns and the tone sandhi patterns to show that the overall analysis works; a challenge that I will pursue in this thesis (see Chapter 5).

4.5 Taiwanese (Zhang & Lai, 2007; Hsiao, 2018)

4.5.1 Taiwanese AABB (Hsiao, 2018)

Taiwanese AABB is exemplified below.

(35) a. *ki33 kuai21*
strange weird
'strange'

ki33 ki33 kuai53 kuai21
strange strange - weird
'very strange'

b. *khin33 sang55*
light loose
'relaxed'

khin33 khin33 sang33 sang55
light light - loose
'very relaxed'

Note that Hsiao (2018) only discusses the base-reduplicant alignment and the reduplicant's shape; there is no mention of the phonological material. Given the form AABB, there are four logically possible alignments between the base and the reduplicant, similar to the Chengdu Mandarin case. Each form corresponds to a specific constraint ranking involving stem-alignment and reduplicant-alignment constraints, defined in (36)–(39).

(36) ALIGNRED-R: For every RED, assign one violation mark for every syllable that intervenes between the right edge of that RED and the right edge of the Wd.

(37) ALIGNRED-L: For every RED, assign one violation mark for every syllable that intervenes between the left edge of that RED and the left edge of the Wd.

(38) AlignStem-R: For every stem, assign one violation mark for every syllable that intervenes between the right edge of that stem and the right edge of the Wd.

(39) AlignStem-L: For every stem, assign one violation mark for every syllable that intervenes between the left edge of that stem and the left edge of the Wd.

The proposed constraint ranking is (40), in which the reduplicant is a suffix.

(40) $A_1\overline{A_2}B_1\overline{B_2}$: AlignStem-L, AlignRed-R ≫ AlignStem-R, AlignRed-L

There is both cross-linguistic and internal evidence supporting this ranking. One cross-linguistic observation is that full reduplication in most Sinitic languages patterns as suffixation. In Pingnan Baihua, for instance, the expressive suffix *kuai* appears in the second syllable position where the reduplicative would be (i.e., A *kuai* BB), suggesting that the reduplicant occupies a suffixal position in the form $A\overline{A}B\overline{B}$. The evidence from within Taiwanese comes from literary Taiwanese, in which the emphatic marker *tsai* appears in the second and fourth syllables of the construction X.X.Y.Y, again pointing to the interpretation that the first and third syllables are the base, and the second and fourth are the reduplicant, located where a regular suffix would be.

To further refine the analysis, a set of Contiguity constraints is added to the alignment ranking, yielding a revised constraint ranking in (41).

(41) Revised (40):
$A_1\overline{A_2}B_1\overline{B_2}$: AlignStem-L, AlignRed-R ≫ AlignStem-R, AlignRed-L, ContigRed, ContigStem

While both Hsiao (2018) and the Chengdu Mandarin paper make use of Contig constraints, they do so in opposite ways. In Chengdu Mandarin, Contig is invoked to explain *why* $A_1\overline{A_2B_1}B_2$ surfaces as optimal, whereas in the Taiwanese case, Contig is ranked lower to allow the $A_1\overline{A_2}B_1\overline{B_2}$ pattern to win, despite its violation of contiguity.

Because AABB in Taiwanese is a case of full reduplication, the focus of this analysis is on the alignment between base and reduplicant, rather than tone sandhi. As such, it is not

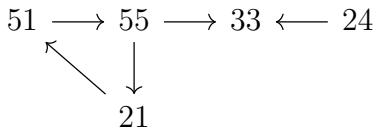
clear what influences, implications, or interactions of the alignment constraints and CONTIG have in other reduplication processes or tone sandhi in Taiwanese.

AABB in Taiwanese bears a closer resemblance to Vietnamese AABB reduplication than to that of Chengdu Mandarin. In both Taiwanese and Vietnamese, the process is full reduplication. However, while RED behaves as a suffix in Taiwanese (i.e., $A_1\overline{A_2}B_1\overline{B_2}$), in Vietnamese, it appears to be a prefix (i.e., $\overline{X}.X.\overline{Y}.Y$).[11]

4.5.2 Taiwanese triplication and tone sandhi (Zhang & Lai, 2007)

Taiwanese tone sandhi, like other Southern Min dialects, is positionally conditioned. Specifically, tones undergo sandhi in non-XP-final positions, while the tone of the XP-final syllable remains unchanged. This positional dependency is illustrated in the tone sandhi pattern in (42).[12]

(42) Taiwanese tone sandhi in non-XP-final positions:



Because tone sandhi applies only in non-final positions, the analysis assumes that in reduplication, the final syllable is the base while the preceding syllables are reduplicants. In a trisyllabic reduplicative form, the second syllable undergoes general tone sandhi. In contrast, the first syllable behaves differently depending on the base tone. It follows the general sandhi pattern only when the base tone is 21 or 51. When the base tone is 55, 33, or 24, however, the first syllable surfaces with a rising tone $\overline{\mathbf{35}}$. This distribution is summarized in (43).

[11] I used XY notation for Vietnamese to avoid confusion with A and B tones.

[12] The five tones considered here are for unchecked syllables, which are syllables that are either open or closed with a sonorant coda.

(43) Tone patterns in Taiwanese reduplication

Monosyllabic base	Single reduplication	Double reduplication[13]
21	**51**-21	**51**-**51**-21
51	**55**-51	**55**-**55**-51
55	**33**-55	**35**-**33**-55
33	**21**-33	**35**-**21**-33
24	**33**-24	**35**-**33**-24

To derive these patterns, the authors propose two main mechanisms: (1) the docking of a floating High (H) tone and (2) the selection of appropriate tonal allomorphs for non-final syllables. The floating H tone is key to understanding the tone on the first syllable of trisyllabic reduplication. The sandhi tone on the second syllable is first determined through allomorph selection, and the tone on the first syllable is then derived by docking a floating H tone to the sandhi tone. If the sandhi tone starts with a H tone (i.e., 51 or 55), the floating H docks to the left edge of the syllable. If the sandhi tone starts with a mid tone (i.e., 33 or 21), the floating H docks to the right edge.

This analysis is supported by a set of ranked constraints.

(44) Realize(Float): A floating tone must be realized in the output.

(45) Align(Float, L, Wd, L) (abbr. Align-L(Float, Wd)): The left (L) edge of a floating tone must be aligned with the left edge of a word (Wd).

(46) Ident-RR(Tone, L) (abbr. Id-RR(T, L)): The left edges of the tones of two reduplicants derived from the same base must be identical.

(47) Ident-RR(Tone) (abbr. Id-RR(T)): The tones of two reduplicants derived from the same base must be identical.

Realize(Float) is undominated since all outputs in the dataset realize the floating H tone. Id-RR(Tone, L) is ranked above Align-L(Float, Wd), favoring outputs in which the left edge of the two reduplicants derived from the same base are identical. Id-RR(Tone) is

[13]Triplication is called *double reduplication* in Zhang and Lai (2007).

ranked lowest because identity between the tones of the two reduplicants is not observed except in cases where the base tone is 21 or 51. This ranking is formalized in (48).

(48) Realize(Float), Id-RR(Tone, L) ≫ Align-L(Float, Wd) ≫ Id-RR(Tone)

The tone sandhi in non-final syllables is treated as an *allomorph selection* process rather than a productive phonological rule, especially since such alternations fail to generalize to novel words (Tsay & Myers, 1996). This is captured by the constraint Appropriate-Allomorph (Allmph), which selects surface tonal allomorphs from the lexicon.

(49) Appropriate-Allomorph (abbr. Allmph): For an existing syllable, select its surface tonal allomorph as follows:

UR	Surface allomorph	
	XP-final	Non-XP-final
21	21	51
51	51	55
55	55	33
33	33	21
24	24	33

The revised ranking is in (50). Realize(Float) and Id-RR(Tone, L) rank above the Allmph constraint, ensuring that (1) no cases where the two reduplicant tones are similar to each other but does not realize the floating H tone in the output, and (2) the floating H being on the right edge of the first syllable when the base has a sandhi tone that does not have an initial H. In turn, Allmph is ranked above Align-L(Float, Wd). For instance, with a base tone 55, an RR sequence like 35-33 and 53-53 are possible outputs, 53-53 does not violate Align-L(Float, Wd) but the supposed output 35-33 does, so Allmph must dominate to rule out cases like 53-53 because while 35-33 violates that once (for the 35 in the first syllable instead of 33), 53-53 violates that twice.

(50) Revised (48):

REALIZE(FLOAT), ID-RR(Tone, L) ≫ ALLMPH ≫ ALIGN-L(Float, Wd) ≫ ID-RR(Tone)

Zhang and Lai (2007) assume a fixed alignment between base and reduplicant and does not examine possible interactions between alignment constraints and the tonal constraints discussed. Comparing this analysis with that of the Chengdu Mandarin, both authors deal with floating H tones but propose different constraint sets. The Chengdu analysis uses *FLOAT and MAXFLOAT to ban unassociated tones and penalize their deletion, respectively. In contrast, Zhang and Lai (2007) uses REALIZE(FLOAT), which stems from Kurisu (2001)'s REALIZE MORPHEME constraint, requiring every underlying morpheme to receive some phonological exponent.

In addition, both Zhang and Lai (2007) and Lin (2015) analyze the pattern of single reduplication differently. In Chengdu Mandarin, since this pattern is productive, Lin (2015) analyzes this as a docking of floating H tone in the reduplicant. As for Taiwanese, since this pattern is *not* productive, Zhang and Lai (2007) proposes the constraint ALLMPH following Tsay and Myers (1996).

Note that although the tone sandhi pattern in single reduplication is not productive, speakers still produce the first-syllable tone in double reduplication presented in (43) even when they get the second-syllable tone that is not presented in (43) or not following the pattern shown in (42) (Zhang & Lai, 2007). There seems to be a split in productivity: one reduplicant (the first syllable) is derived via a productive process involving the docking of the floating H tone, while the other reduplicant (the second syllable) reflects a lexicalized, non-productive process of allomorphy selection.

4.6 Interim summary

This section summarizes how triplication and AABB patterns have been described and analyzed both cross-linguistically and within Vietnamese. While triplication and reduplication mechanisms more broadly have been extensively discussed in Vietnamese literature,

three representative works are selected here: the descriptive accounts of P.-P. Nguyen (1998) and Hoang et al. (2008), and the formal analysis of Vu (2007). Two salient features of triplication are the apparent identity avoidance and a tone pattern that diverges from expectations. However, one possible approach that has been largely overlooked is the idea that triplication may not involve multiple rounds of base copying. Instead, it could involve *allomorph selection*. From this perspective, what looks like reduplication may reflect the presence of *two separate morphemes*, each contributing to the phonological representation and interpretation of the overall reduplicative form. This alternative analysis will be developed further in Section 5.8.

Since the analysis of triplication and AABB reduplication will make reference to the same processes in other languages, a review of these processes in Chengdu Mandarin and Taiwanese were discussed as well (Zhang & Lai, 2007; Lin, 2015; Hsiao, 2018). One noticeable observation is that the analyses provided by these works are either of alignment or tone sandhi only, or both but *separately*. Hence, they do not show how the two analyses might interact with each other. The analysis in this thesis will attempt to close that gap in a similar analysis for Vietnamese reduplication; it is crucial to show how the two patterns of alignment and tone interact with each other for the sake of a uniform analysis of reduplication. Such a holistic approach contains possible analyses for tones; for instance, one cannot say something is a default tone if it has incorrect consequences on other tone patterns.

CHAPTER 5

ANALYSIS

Although the core of this study centers on tone patterns and base-reduplicant alignment in triplication and AABB reduplication, re-analyzing other productive reduplication patterns in Vietnamese, especially those previously treated in OT (Vu, 2007), offers valuable insight toward a unified OT account. Such an approach makes it possible to evaluate how tone interacts with alignment constraints across the full range of productive reduplication, contributing to a broader understanding of Vietnamese morphophonology.

This chapter is arranged as follows. The relevant theoretical models that the analysis utilizes are briefly discussed in Section 5.1 and Section 5.2. From Section 5.4 to Section 5.8, the analysis of each reduplication process — attenuation, pejorative, intensification, exclamation, and triplication — is presented.

5.1 Correspondence Theory

This analysis is presented in the *Optimality Theory* (OT) framework, in which languages differ by how constraints are ranked against each other (Prince & Smolensky, 2002). Under this theoretical framework, given some input, an infinite set of output candidates are generated by the function GEN and evaluated by the function EVAL for well-formedness against a set of ranked and violable constraints. The optimal candidate is selected as the one that violates the lowest-ranked constraints and violates the dominant constraints the least number of times.

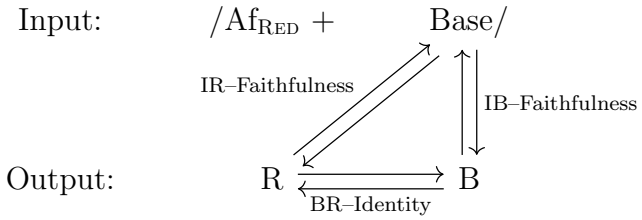
The model in OT that is of interest here is Correspondence Theory (McCarthy & Prince, 1995) which involves correspondence relation between stem and base, between base and reduplicant, and between stem and reduplicant. The correspondence relation is defined in (1).

(1) **Correspondence** (McCarthy & Prince, 1995): Given two strings S_1 and S_2 , *correspondence* is a relation $\mathfrak{R}$ between the elements of S_1 and those of S_2. Elements

$\alpha \in S_1$ and $\beta \in S_2$ are *correspondents* of one another when $\alpha \,\mathfrak{R}\, \beta$.

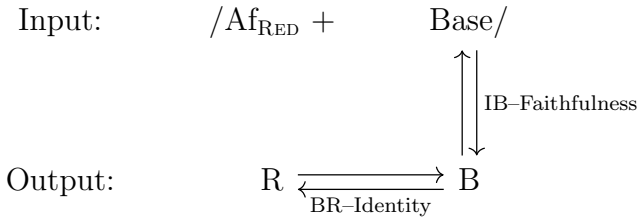
The system of relations is schematized as follows,

(2) Full model



For this analysis presented in this chapter, the basic version of the model above will be used instead since it is sufficient enough. The basic model is shown below.

(3) Basic model



Constraints must govern the correspondence and identity of correspondent elements. There are separate constraints for each I(nput)-O(utput) and B(ase)-R(eduplicant) correspondence relation. The following are three of the constraint families that will play a role in the analysis with its corresponding domain-specific instances:

(4) MAX constraint family: Every segment of string S_1 has a correspondent in string S_2.

   a. MAX-IO: Every segment of the input has a correspondent in the output.

   b. MAX-BR: Every segment of the base has a correspondent in the reduplicant.

(5) DEP constraint family: Every segment of string S_2 has a correspondent in string S_1.

   a. DEP-IO: Every segment of the output has a correspondent in the input.

   b. DEP-BR: Every segment of the reduplicant has a correspondent in the base.

(6) IDENT constraint family: S_2 correspondents of S_1's [γF] segment are also [γF].

   a. IDENT-IO: Output correspondents of an input [γF] segment are also [γF].

   b. IDENT-BR: Reduplicant correspondents of the base [γF] segment are also [γF].

Among these constraint families, MAX will be used most intensively. One reason is that the difference the input and the output, mainly involves *no deletion* of segments or features that are specified in the input for either the base or the reduplicant.

5.2 Prosodic hierarchy

Following Vu (2007), the prosodic hierarchy will be used in this analysis for the description of stress patterns and syntactic separability in Vietnamese. In the prosodic hierarchy, sentences can be a structure such that its categories can be organized hierarchically such that in each domain, various phonological and morphological processes can take place (Selkirk, 1978; Hayes, 1989, et seq.).

Prosodic hierarchy	**Category**
Utt	Utterance
IP	Intonational phrase
ϕ	Phonological phrase
ω	Prosodic word
Ft	Foot
σ	Syllable
μ	Mora

5.3 Phonotactics

I posit the constraints listed below, along with MAX-IO, to be undominated in the language since no inputs nor outputs violate them. The tone-coda co-occurrence restriction is not the main focus of this analysis; thus, an undominated arbitrary markedness constraint, PHONO, will be posited. The unpacking of this constraint into more motivated constraints is left to future work. Because no outputs in the language violate the phonotactic constraint, it is assumed to be dominant. This ensures that any sequence of obstruent coda with A or C tones is barred from the inventory of the language as a whole, and hence will never have a surface representation.

Following Vu (2007), I assume that every foot is binary and hence FtBin (σ), where each foot must consist of two syllables, in Vietnamese. The constraint's definition is given in (7).

(7) FtBin (σ): Feet are binary under syllabic analysis. (Prince and Smolensky, 2002).

This is enforced by ParseSyll, which requires that every syllable be parsed into a foot. Stress assignment is handled by Align-R(Ft, H(Ft)), which demands that the right edge of a foot align with that of its head, ensuring that stress consistently falls on the rightmost syllable.

(8) ParseSyll: All syllable σ must be parsed by a foot Ft.

(9) Align-R(Ft, H(Ft)): For every foot Ft, there's a head foot H(Ft) such that the right edge of Ft coincides with the right edge of H(Ft).

To prevent empty syllables in outputs, a Fill constraint is assumed to be undominated. This constraint penalizes any syllabic position that is not filled by segmental material. In the context of Base-Reduplicant Correspondence Theory (BRCT), the reduplicant Red is a morpheme that is phonologically empty in the input (McCarthy & Prince, 1994). Candidates that insert the reduplicant without filling it will violate Fill and are eliminated. Note that satisfying Fill in the output does *not* violate Max-IO, since the reduplicant is phonologically empty in the input and thus has no input correspondents. It is also assumed that the shape of the reduplicant is minimally a syllable. Previous work in OT analyses of reduplication has proposed various ways to derive this through constraint interactions.[1] However, since this issue is not within the scope of the present analysis, I will assume a minimal syllabic shape for the reduplicant.

(10) Fill: Syllable positions are filled with segmental elements. (Prince and Smolensky, 2002)

[1] One consequence of template derivation is that a Realize Morpheme constraint (Kurisu, 2001) could be used in place of Fill, but since no derivation was attempted, Fill is used instead.

5.4 Attenuation

In attenuation, there are two main patterns: *full* reduplication, where the reduplicant is an exact copy of the base, and *partial* reduplication, where the reduplicant exhibits a fixed flat tone (i.e., A tone) while the remaining segmental material is copied from the base. Note that attenuation reduplicatives are *not* subject to the syntactic separability effect. Thus, an assumption is made, which is attenuation reduplicatives are of a Pr(osodic) W(or)d (ω) and therefore also a foot. Each of the base and reduplicant is then a syllable in size. The following alignment constraints are adopted to regulate the positions of both the base and reduplicant within the PrWd. Ultimately, we need specific ALIGN for different reduplicant allomorphs, but in the interest of brevity, the constraints are presented as such.

(11) ALIGN-L(Base, ω) (BASE-L(ω)): For every base, there's a PrWd ω such that the left edge of the of the base coincides with the left edge of ω.

(12) ALIGN-R(Base, ω) (BASE-R(ω)): For every base, there's a PrWd ω such that the right edge of the of the base coincides with the right edge of ω.

(13) ALIGN-L(RED, ω) (RED-L(ω)): For every RED, there's an ω such that the left edge of RED coincides with the left edge of the PrWd.

(14) ALIGN-R(RED, ω) (RED-R(ω)): For every RED, there's an ω such that the right edge of RED coincides with the right edge of the PrWd.

In attenuation, the reduplicant surfaces as a prefix. This prefixal position satisfies ALIGN-L(RED, ω) and ALIGN-R(Base, ω) but violates ALIGN-L(Base, ω), as the base is no longer aligned with the left edge of the PrWwd due to the insertion of the reduplicant. The alignment ranking that captures this is presented below with the corresponding tableau. Along with the undominated constraints that govern Vietnamese phonotactics, these alignment constraints are dominant. For simplicity, only relevant alignment constraints present in the tableau.

(15) ALIGN-L(RED, ω), ALIGN-R(Base, ω) $\gg$ ALIGN-R(RED, ω), ALIGN-L(Base, ω)

B + R	RED-L(ω)	BASE-R(ω)	RED-R(ω)	BASE-L(ω)
a. ☞ [(σ_R.σ_B)]			*	*
b. [(σ_B.σ_R)]	*!	*!		

5.4.1 Full reduplication

In full reduplication, the reduplicant is a total copy of the base. Hence, both MAX-IO and MAX-BR are satisfied; the input base is preserved in the output, and there is no deletion of segments nor mismatch between the base and its reduplicant. Markedness constraints (M) are dominated and therefore do not influence the choice of output. The optimal candidate is the one that fully satisfies faithfulness to both input and base, provided it also meets the undominated phonotactic constraint and is not empty segmentally. Thus, the desired ranking is in (16).

(16) PHONO, FILL, ALIGN-L(RED, ω), ALIGN-R(Ft, H(Ft)), ALIGN-R(Base, ω), PARSESYLL, FTBIN, MAX-IO, MAX-BR $\gg$ M, ALIGN-R(RED, ω), ALIGN-L(Base, ω)

Consider the example tableau below for the input /saːŋB1/ 'bright'. For simplicity and relevance of empty the reduplicant, of the dominant constraints besides MAX-IO and MAX-BR, only FILL will be in the tableau. Candidate (b) and (c) are ruled out because they each violates MAX-BR and MAX-IO respectively. Candidate (d) is ruled out because it violates the undominated FILL for having an empty syllable. This leaves candidate (a) to be the optimal output.

(17) Input: /saːŋB1/ 'bright'

R + saːŋB1		Fill	Max-IO	Max-BR
a.	☞ saːŋB1.saːŋB1			
b.	saːŋA1.saːŋB1			∗!
c.	saːŋA1.saːŋA1		∗!	
d.	R.saːŋB1	∗!		

5.4.2 Partial reduplication

Since the reduplicant's tone has a fixed A contour and copies the base tone register, we can either approach this as the fixed tone contour arising through different constraint rankings (i.e., A tone is phonologically derived) or it being lexically specified in the reduplicant input (i.e., an allomorph of the reduplicant).

A set of markedness constraints can be imposed to opt for the flat tone to occur in the reduplicant at the expense of an unfaithful mapping between the base and reduplicant. For the purpose of the analysis, a modified version of NoContour in Pulleybank (1997), which prohibits more than one tone to be linked to a vowel, is proposed. Recall from Section 2.1 that in their tone structure model, flat tones only have *one* point linked to its contour node. Thus, this constraint will rule out non-flat tone, which has more than one point under their contour node.

(18) NoContour: No more than one point may be linked to the contour node.

The proposed ranking, shown in (19), will follow the model of *The Emergence of The Unmarked* (TETU), in which Faith-IO ≫ M ≫ Faith-BR (McCarthy & Prince, 1994). Under this schema, the inactive markedness constraint M (due to being dominated by Max-IO) becomes active in BR mappings where IO faithfulness is irrelevant. The ranking points out that any tone and not necessarily flat tones is permitted in the language, but because

NoContour ≫ Max-BR, the reduplicant will not copy the tone from the base and has to have a flat tone.

(19) Max-IO ≫ NoContour ≫ Max-BR

Similarly, another way to approach fixed segmentism here is to consider the tone contour A and tone register α being specified in the reduplicant as a requirement of the lexicon. If that's the case, the input will be something like $\text{Base.red}_{\text{A}\alpha}$ in which Base represents the phonological material of the base and Aα subscript represents the specified segment in the reduplicant. With the reduplicant being specified with a tone contour and tone register already, the only thing that is left to do is filling in any of the empty positions, in this case, onset, nucleus, and coda and thus the ranking below.

(20) Fill, Max-IO ≫ Max-BR

Note that this is the same ranking for full reduplication in (16). Previously, Phono and Fill have been established to be undominated in the constraint ranking to (1) rule out any input or output that violates the phonotactic constraint on tone and coda, and (2) no empty syllable can have a surface representation. On the contrary to the case of fixed tone through constraint ranking, NoContour is not needed here. Max-IO dominates Max-BR to rule out any candidate that alters the phonological content of the base or the reduplicant as an effort to obey Max-BR.

Consider the input /saːkB1/ 'to be close to' in the tableau below. Candidate (c) is eliminated for violating the highly ranked Fill. Between candidates (a) and (b), (a) is the winner because (b) violates Max-IO for deleting the flat contour to preserve the identity relation between the base and reduplicant. Candidate (a) is the optimal output at the expense of deleting the non-flat contour in the reduplicant.

(21) Input: /saːkB1/ ‘to be close to’

$R_{A\alpha}$ + saːkB1		Fill	Max-IO	Max-BR
a.	☞ saːŋA1.saːkB1			*
b.	saːkB1.saːkB1		*!	
c.	$R_{A\alpha}$.saːkB1	*!		*

Comparing the two analytical options, the phonological-derivation approach requires an additional markedness constraint, NoContour, which complicates the constraint hierarchy. In contrast, the allomorphy case relies on a simpler ranking that only necessitates tone register identity. This analytical choice becomes particularly significant when we consider other reduplication processes — specifically exclamation/AABB forms — which only permit the full reduplication ranking shown in (16). For the purpose of developing a unified OT analysis and avoiding conflicting constraint rankings within the same prosodic domain, I will adopt the allomorphy approach for the attenuation process, as presented below. This decision eliminates the need for multiple competing constraint hierarchies within Vietnamese, providing a more coherent overall account of the language’s reduplication.

(22) *Revised* (16)

Phono, Fill, Align-L(Red, ω), Align-R(Ft, H(Ft)), Align-R(Base, ω), ParseSyll, FtBin, Max-IO $\gg$ Max-BR, Align-R(Red, ω), Align-L(Base, ω)

5.5 Pejorative

Recall from Section 3.3 that the reduplicant in the pejorative process is either [-iəkB1] or [-iəkB2]. The process disallows any input that has the same rhyme [-iək] as the reduplicant’s. In Southern Vietnamese, there is a diphthong reduction process in which the rime /-iək/ surfaces as [-iːk]. SVN speakers tend to omit the second vowel and lengthen the first as a compensatory strategy.[2] This reduction can be explained within the framework of OT

[2]This process of diphthong reduction is observed in all three diphthongs [iə, uə, uə] in SVN.

using a universally ranked family of markedness constraints on place features: *Pl-Dor, *Pl-Lab ≫ *Pl-Cor, where coronals are the least marked compared to dorsals and labials (Lombardi, 1997; Prince & Smolensky, 2002). The nucleus [-iː] then emerges in the output due to a ranking *Pl-Dor, *Pl-Lab ≫ *Pl-Cor ≫ Max-BR. However, this still does not explain why the coda surfaces as [-k], a dorsal segment, despite dorsals being the most marked under this ranking.

Since the rime contains a marked structure that cannot be explained by phonological constraints alone, it is assumed to be specified lexically. In other words, the specification [-iək] in $R_{iək}$ in pejorative reduplication is lexically provided rather than derivable solely from phonology.

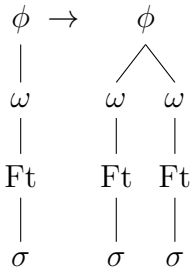
Regarding the tone, there are three possible underlying representations of the reduplicant that can be considered: full specification /$R_{iək}$ B1/, partial specification /$R_{iək}$ Bα/, and no specification /$R_{iək}$/. Given the fixed B1, the full specification /$R_{iək}$ B1/ needs a markedness constraint on the reduplicant's tone to dominate Max-IO in cases where the base with a low tone (i.e., tone register = 2) wants the reduplicant to have a low tone too at the expense of the output being unfaithful to the input. However, a consequence of this would be that if M ≫ Max-IO, Max-BR, the segment or structure militated against by M (in this case, a high tone) is barred from the inventory of the whole language.

Another possibility is the partial specification of the tone contour /$R_{iək}$ Bα/. However, this appears redundant because phonotactic constraints Phono, Max-BR, and possibly other markedness constraints already favor B1 or B2 as appropriate.

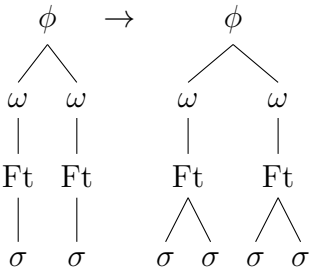
The last possibility is no specification /$R_{iək}$/. In this case, phonotactic constraints generate B tones, and Max-BR allows either B1 or B2 depending on the tone of the base. When the tone register of the reduplicant matches that of the base, Max-BR will ensure that. In contrast, when the tone freely alternates between B1 and B2, other tone-related markedness constraints might dominate Max-BR to yield the correct output, as will be discussed later. This is the best and least problematic option.

Turning to alignment constraints, the reduplicant $R_{i\text{ə}k}$ must be properly aligned with the phonological phrase (ϕ) (Vu, 2007). This is motivated by the observation that there can be additional syntactic constituents separating the base from the reduplicant in the pejorative process (see Section 3.3 for examples).

Monosyllabic base X (14):



Polysyllabic base X.Y (15):



The constraint Align-R($R_{i\text{ə}k}$, ϕ) ($R_{i\text{ə}k}$-R(ϕ)) ensures that the right edge of $R_{i\text{ə}k}$ coincides with the right edge of the phonological phrase. This is the opposite of attenuation reduplication, in which the reduplicant is prefixed. Furthermore, inserting $R_{i\text{ə}k}$ as a suffix violates the base counterpart of the alignment constraint, Align-R(Base, ϕ) (Base-R(ϕ)).

A second alignment constraint, Align-L($R_{i\text{ə}k}$, ϕ) ($R_{i\text{ə}k}$-L(ϕ)), is the leftward equivalent of the one above. It is always violated in pejorative reduplication because $R_{i\text{ə}k}$ always follows the base, and hence its left edge will never align with the left edge of the phonological phrase.

For the stem, which is assumed to be a PrWd, two alignment constraints are posit: Align-L(Base, ϕ) (Base-L(ϕ)) and Align-R(Base, ϕ) (Base-R(ϕ)), ensuring that both edges of the base align with the phonological phrase. Because $R_{i\text{ə}k}$ follows the base, its right edge must align with that of ϕ, which means Align-R($R_{i\text{ə}k}$, ϕ) and Align-L(Base, ϕ) are satisfied, while Align-L($R_{i\text{ə}k}$, ϕ) and Align-R(Base, ϕ) are violated.

This results in the following alignment ranking,

(23) FtBin, Align-R($R_{i\text{ə}k}$, ϕ), Align-L(Base, ϕ) $\gg$ Align-L($R_{i\text{ə}k}$, ϕ), Align-R(Base, ϕ)

The lower-ranked constraints Align-L($R_{i\text{ə}k}$, ϕ) and Align-R(Base, ϕ) are not ranked against Fill but should dominate Max-IO altogether, similar to the attenuation process.

In tableaux for attenuation reduplication, FtBin (and Align-R(Ft, H(Ft))) were not shown because the reduplicant is always a disyllabic PrWd, and hence these are dominant. However, in the case of pejorative reduplication, some inputs have stems composed of two PrWds (e.g., coordinate compounds). Hence, FtBin is necessary in this analysis for completeness as to eliminate candidates that (1) retain the prosodic structure of the base or (2) have a reduplicant composed of two separate PrWds.

(24) Alignment tableau for Base and $R_{iək}$

B + $R_{iək}$		FtBin	$R_{iək}$-R(ϕ)	Base-L(ϕ)	$R_{iək}$-L(ϕ)	Base-R(ϕ)
a. ☞	$[(\sigma.\sigma)]_B$ $[(\sigma.\sigma)]_R$				*	*
b.	$[(\sigma)]_B$ $[(\sigma)]_B$ $[(\sigma.\sigma)]_R$	*!*		*!	*	**
c.	$[(\sigma.\sigma)]_B$ $[(\sigma)]_R$ $[(\sigma)]_R$	*!*	*!		**	*

In pejorative reduplication, because the rime [-iək] is lexically specified, the only parts of the reduplicant that need to be filled are the onset and tone positions. The onset is copied faithfully from the base. The tone, however, may mismatch if the base has tone A or C, since these are phonotactically incompatible with the coda [-k]. Thus, the tone of the reduplicant must be adjusted, and the following ranking emerges,

(25) Phono, Fill ≫ Max-IO ≫ Max-BR.

Since there is no markedness constraint between Max-IO and Max-BR, the most optimal output will be the one that preserves the segmental content of the input and only changes the tone contour if necessary due to phonotactic restrictions. Novel data also suggest that the reduplicant tends to adopt the tone register of the base. In such cases, a markedness constraint like *L, which bans low tones, must dominate Max-BR in order for B1 to surface when the base has a low tone.

(26) *L: Low tones are prohibited. (Bickmore, 2000)

Thus, the revised ranking is as follows,

(27) FILL, MAX-IO ≫ *L ≫ MAX-BR

The ranking in (27) only emerges when the base has a low tone. However, this ranking will be in conflict with the ranking in (25) because there cannot be two conflicting rankings for one morpheme. It follows then the ranking between *L and MAX-BR is variable; there is no strict order between them, allowing either constraint to dominate the other in different bases. However, this is not within the scope of the analysis, so I will leave this for future research. In the meantime, the *no-alternation* case is assumed for this analysis. The full ranking is presented below.

(28) PHONO, FILL, ALIGN-R(Ft, H(Ft)), PARSESYLL, FTBIN, MAX-IO, ALIGN-R($R_{iək}$, ϕ), ALIGN-L(Base, ϕ) ≫ ALIGN-L($R_{iək}$, ϕ), ALIGN-R(Base, ϕ), MAX-BR

The ranking works as follows. Consider the input /lanA2/ 'mild' in the tableau below. Candidate (c) and (d) are eliminated because each violates FILL and MAX-IO. Between candidates (a) and (b), candidate (a) is the optimal output because it has less violation in MAX-BR than candidate (b) does.

(29) Input: /lanA2/ 'mild'

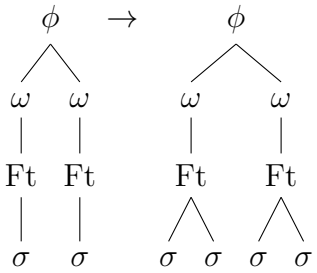
lanA2 + $R_{iək}$	FILL	MAX-IO	MAX-BR
a. ☞ lanA2.liəkB2			***
b. lanA2.liəkB1			****!
c. lanA2.lanA2		*!**	
d. lanA2.$R_{iək}$	*!		***

5.6 Intensification

Recall from Section 3.4 that the reduplicant in the [-aː] reduplication always has a [-aː] rhyme and a flat tone with its register copied from that of the base. Similar to the pejorative process, this also avoids inputs that have the same [-aː] rhyme as the reduplicant

does. Regarding the alignment of the base and reduplicant, [-aː] reduplication has a prefixal reduplicant similar to that in the attenuation process. One noticeable difference is that since there is separability effect observed in the reduplicative, each of the base and the reduplicant is a Prwd and hence also a Ft. As a result, the whole reduplicative is a phonological phrase, mirroring the prosodic structure of the reduplicative from the pejorative process, illustrated below.

Polysyllabic base X.Y (19):



This results in the alignment ranking in (30). It is similar to that of the pejorative reduplication as both share a similar prosodic domain governed by the same alignment constraints. The only difference is that in this case, the left-edge Align for the reduplicant and right-edge counterpart for the base dominate the right-edge Align of the reduplicant and left-edge Align of the base, which makes the reduplicant a prefix, similar to the reduplicant of the attenuation process.

(30) FtBin, Align-L($R_{aː}$, ϕ), Align-R(Base, ϕ) ≫ Align-R($R_{iək}$, ϕ), Align-L(Base, ϕ)

Regarding the phonological material of the reduplicant, there are two possible cases to consider: /R-aː/ or /R-aːAα/. In the first case, the tone in the reduplicant is phonologically derived. The possible ranking is Max-IO ≫ NoContour ≫ Max-BR.

The markedness constraint NoContour motivates the choice of an A tone in the reduplicant, while Max-BR ensures that the tone register is preserved. This latter ranking resembles that of the TETU model, which as a result inferring that the flat tone is unmarked.

This is also parallel to the partial specification in the reduplicant of the attenuation and pejorative processes.

In the second case, the reduplicant is specified with both the rhyme and the tone contour [-aːA]; the tone register is derived phonologically. Here, a simple ranking of MAX-IO ≫ MAX-BR is sufficient to derive the correct output, selecting the tone register α that matches the base. This is similar to the ranking for full specification of the reduplicant in the attenuation and pejorative reduplication in (16) and (25) respectively.

Comparing the two options, the /R-aː/ case requires the addition of the markedness constraint NOCONTOUR, which introduces additional candidates and complexity. In contrast, the /R-aːAα/ case relies on a simpler ranking that only necessitates tone register identity and involves fewer competing candidates. More importantly, the former ranking with NOCONTOUR would create a direct conflict with the rankings established for full reduplication processes (attenuation, pejorative, and later, AABB) within the same prosodic domain, as these processes do not require a flat tone in the reduplicant. Since a unified analysis should avoid contradictory constraint rankings within the same language system, this analysis adopts the /R-aːAα/ case, which maintains consistency with the constraint hierarchy established for other reduplicative processes in Vietnamese.

In the interest of space, I will not provide example tableaux for the constraint ranking of the intensification process in (31). However, since the rhyme and the tone contour are specified in the input, the reduplicant only needs to copy the onset and the tone register of its base. One thing to note is that, given the form $\overline{X_1 Y_1} X_2 Y_2$, the specification of [-aːAα] is for the second syllable of the reduplicant, $\overline{Y_1}$; $\overline{X_1}$ remains a total copy of X_2.

(31) PHONO, FILL, ALIGN-R(Ft, H(Ft)), PARSESYLL, FTBIN, MAX-IO, ALIGN-L($R_{aː}$, ϕ), ALIGN-R(Base, ϕ) ≫ ALIGN-R($R_{iək}$, ϕ), ALIGN-L(Base, ϕ), MAX-BR

5.7 Exclamation (AABB)

The alignment analysis for both the base and the reduplicant, as well as the positioning of the reduplicant within the output of the exclamation process, follows the work of Lin

(2015) and Hsiao (2018). Both of these works analyze reduplication processes in Chengdu and Taiwanese, respectively, that involve full reduplication of the form AABB, which is structurally similar to the reduplication seen in this exclamation process.

This type of reduplication involves total copying, meaning the reduplicant is a faithful reproduction of the base. Consequently, the phonological exponent of the reduplicant does not pose significant challenges, as long as the faithfulness constraints MAX-IO and MAX-BR dominate all markedness constraints, except those necessary for phonotactics and alignment. Thus, the ranking will be similar to that of full reduplication and specification in attenuation, pejorative, and intensification.

Now, let us turn to alignment. Two constraints are adopted from Hsiao (2018) that govern the positioning of the base within the prosodic word.[3] First, ALIGN-L(Base, ω) ensures that the left edge of the base aligns with the left edge of the prosodic word. Second, ALIGN-R(Base, ω) requires that the right edge of the base coincide with the right edge of the PrWd. Similarly, alignment constraints for the reduplicant are posited: ALIGN-L(RED, ω) and ALIGN-R(RED, ω) require that the reduplicant align with the prosodic word at its respective edges (Hsiao, 2018).

To further capture the mapping between the reduplicant and the base, I also include contiguity constraints. Following Lin (2015) and Hsiao (2018), CONTIGBASE and CONTIGRED are adopted, as defined below,

(32) CONTIGBASE: Let base = $s_1 s_2 \ldots s_n$ and $\omega = w_1 w_2 \ldots w_m$
Assign one violation mark (*) for every w_y if there's no s_x where $s_x \mathfrak{R} w_y$ (Golston, 1996).

(33) CONTIGRED: Let RED = $r_1 r_2 \ldots r_n$ and $\omega = w_1 w_2 \ldots w_m$
Assign one violation mark (*) for every w_y if there's no r_x where $r_x \mathfrak{R} w_y$ (McCarthy & Prince, 1995).

[3]Hsiao (2018) uses *stem* instead of base in the constraint; for the purpose of the analysis, I change all mentions of "stem" to "base" in the alignment constraints here.

In both attenuation and pejorative reduplication, the base and reduplicant appear to form a contiguous string. However, in the current process, since we do not yet know how the reduplicant is arranged with respect to the base, there is the possibility that contiguity is disrupted, as is potentially the case in candidate outputs like (1), (3), and (6).

This uncertainty arises because we cannot definitively determine which syllables are part of the base and which are part of the reduplicant. Thus, there are four possible arrangements of the reduplicant within the reduplicative, each corresponding to a distinct ranking of alignment and contiguity constraints:

1. X.$\overline{\text{X}}$.$\overline{\text{Y}}$.Y

   Align-L(Base, ω), Align-R(Base, ω), ContigRed $\gg$ Align-L(Red, ω), Align-R(Red, ω), ContigBase

   The reduplicant $\overline{X}.\overline{Y}$ infixes between the base X.Y.

2. $\overline{\text{X}}$.X.Y.$\overline{\text{Y}}$

   Align-L(Red, ω), Align-R(Red, ω), ContigBase $\gg$ Align-L(Base, ω), Align-R(Base, ω), ContigRed

   The reduplicant is split into two, the 1st half $\overline{X}$ is a prefix, and the 2nd half $\overline{Y}$ is a suffix.

3. X.$\overline{\text{X}}$.Y.$\overline{\text{Y}}$

   Align-L(Base, ω), Align-R(Red, ω) $\gg$ Align-R(Base, ω), Align-L(Red, ω), ContigBase, ContigRed

   Both base and reduplicant are split — the 1st half $\overline{X}$ of reduplicant follows its base X, and the 2nd half $\overline{Y}$ of the reduplicant follows its base Y.

4. $\overline{\text{X}}$.X.$\overline{\text{Y}}$.Y

   Align-R(Base, ω), Align-L(Red, ω) $\gg$ Align-L(Base, ω), Align-R(Red, ω), ContigBase, ContigRed

   Both base and reduplicant are split — the 1st half $\overline{X}$ of reduplicant precedes its base X, and the 2nd half $\overline{Y}$ of the reduplicant precedes its base Y.

In all of the possible rankings, two contiguity constraints ContigBase and ContigRed are always dominated because either the reduplicant is split, the base is split, or both are split in the possible arrangements of B and R. For the purpose of this analysis, I assume that (4) is the optimal arrangement as there is no empirical evidence at this point to definitively identify which element is the reduplicant in AABB forms, since both copies are segmentally identical. This configuration reflects full reduplication of each syllable in the base, consistent with the general Vietnamese pattern where the reduplicant precedes the base (cf. attenuation processes and Hoang (1985, 1998) and Cao (2007)).

Accordingly, the complete ranking for the AABB reduplication pattern is as follows,

(34) Phono, Fill, Align-R(Base, ω), Align-L(Red, ω) $\gg$ Align-L(Base, ω), Align-R(Red, ω), ContigBase, ContigRed $\gg$ Max-IO, Max-BR

In this ranking, each reduplicant is a faithful copy of its base, provided that the optimal output satisfies (1) the undominated alignment constraints, (2) the requirement to fill in phonological material (e.g., onset, tone contour, and tone register), (3) faithfulness to the segmental content of the input, and (4) base-reduplicant identity.

The addition of the constraints ContigBase and ContigRed to the ranking creates an important distinction between AABB reduplication and previously discussed processes. In AABB forms, these contiguity constraints must be dominated by alignment constraints, allowing the base or reduplicant to be split. In contrast, for all previously discussed processes (attenuation, pejorative, and intensification), both the base and reduplicant always form contiguous strings. Therefore, in those cases, ContigBase and ContigRed can be undominated without affecting the outputs, as shown in the revised ranking (35). This ranking difference captures the structural distinction between AABB reduplication and other reduplication types, while maintaining consistency in the overall constraint ranking.

(35) Revised (22), with alignment and contiguity constraints added
Phono, Fill, Align-L(Red, ω), Align-R(Ft, H(Ft)), Align-R(Base, ω), ParseSyll, FtBin, Max-IO, ContigBase, ContigRed $\gg$ Max-BR, Align-R(Red,

ω), ALIGN-L(Base, ω)

5.8 Triplication

Recall from Section 3.5 that Vietnamese triplication typically follows specific patterns based on the tone of the base syllable. The general pattern is BASE.R_{A2}.$R_{A1,\ B2}$[4], with the base normally being the first syllable (P.-P. Nguyen, 1998; Hoang et al., 2008). However, several systematic variations occur:

1. When the base has tone B1: The pattern is BASE.R_{A2}.R_{B2}, where R_{B2} has tone B2.
2. When the base has tone A1: The pattern is R_{C1}.R_{A2}.BASE, where R_{C1} has tone C1 (note the reversed order).
3. For all other valid base tones: The pattern is BASE.R_{A2}.R_{A1}, where R_{A1} has tone A1.

Importantly, this process does not permit bases with tone A2. There are two distinct problems regarding triplication — formation and allomorphy selection, which were discussed in Section 3.5. They can be subsumed under one analysis: *allomorphy selection*.

5.8.1 Two separate morphemes

Another hypothesis that has been overlooked is the possibility of having *two separate morphemes* that make up triplication. Consequently, this makes triplication different from simply multiple applications of one reduplication process. If this is the case, the tone and avoidance patterns discussed in Section 3.5 can be formally captured. Crucially, then, Vietnamese triplication is a *morphological*, such that two distinct reduplicative morphemes are concatenated, each with its own lexical entry. Phonology's role is strictly enforcement: high-ranking ALIGN and faithfulness constraints keep each chosen allomorph at its lexically specified edge and tone, but do not decide those properties themselves.

The main evidence for two separate reduplicants is the existence of trisyllabic reduplicatives with corresponding disyllabic forms. Recall (5), repeated below. There are two main differences between (36a) and (36b): (1) their meaning, in which (36a) conveys an emphatic

[4] $R_{A1,\ B2}$ = reduplicant in triplication with A1 or B2

meaning of (36b), and (2) the addition of [saːŋA2] from (36b) to (36a). I suggest that this [saːŋA2] element is a *reduplicant with a fixed A2* that contributes the intensification. The reason this reduplicant consistently bears an A2 tone is due to it being specified in the lexicon. This fixed-tone behavior is attested in Vietnamese, as seen in other examples in Hoang et al. (2008, pp. 78 & 90).

(36) Repeating (5)

a. [saːkB1 saːŋA2 saːkB2]
close.to RED RED
'very close to (EMPH)'

b. [saːkB1 saːkB2]
close.to RED
'very close to'

Having established that the intensification effect arises from the addition of a fixed-A2 reduplicant, consider a similar pattern in (27), repeated below. Unlike the previous example, the disyllabic form in (37a) is illicit, and only the trisyllabic form in (37b) yields the intended intensification. This implies that in such cases, two separate reduplicants are at work: one is the fixed-A2 reduplicant R_{A2} from earlier examples, and the other is an abstract reduplicant RED.[5]

(37) a. * [nɤwC1 **nɤwA1**]
RED brown
* [**nɤwA1** nɤwC1]
brown RED
Intended: 'very brown'

b. [nɤwC1 nɤwA2 **nɤwA1**]
RED R_{A2} brown
'very brown'

(Novel data)

If there are indeed two reduplicants at play, why do we not observe forms where the base combines with R_{A2} alone, or with $R_{A1,\ B2,\ C1}$ alone? A possible explanation is that R_{A2} can only attach to a specific type of base, one that already includes a reduplicative morpheme, as seen in (37b), (36a), and (28a)–(b). In this view, R_{A2} requires a base that already possesses a reduplicative feature, which $R_{A1,\ B2,\ C1}$ supplies as in (37b) and (36a), or a reduplicant supplies as in (28a)–(b).

This selection restriction aligns with Fabb (1988), who shows that suffixes often have constraints on which bases or affixes they can attach to. For instance, in English, the suffix

[5]The form is irrelevant for now but will be discussed later.

-ary selects the suffix *-ion* to create a compound suffix *-ion-ary* that will then select a word to attach to. Since *-ary* is bounded by *-ion*, there are no words in English that has only *-ary* (e.g., *revolution* and *revolutionary* but **revolu-nary*). A consequence of this is that the selectional properties of *-ion* will transfer to *-ion-ary* and hence both apply to the same type of words (i.e., verbs) (Fabb, 1988). Coming back to Vietnamese, we do not observe forms where the base combines with $R_{A1,\ B2,\ C1}$ alone because it is bound by R_{A2}. In return, R_{A2} combines with the base satisfies the selectional property of $R_{A1,\ B2,\ C1}$ that it only selects bases that are disyllabic and already has a reduplicative morpheme.

In conclusion, triplication is not merely the result of iterative reduplication applied to a monosyllabic base. Instead, I suggest that there are two distinct structural patterns: (1) a monosyllabic base accompanied by two distinct reduplicants — $R_{A1,\ B2,\ C1}$ and R_{A2}— and (2) a disyllabic lexicalized base combined with the intensifying morpheme R_{A2}. In the first pattern, intensification results from the coordinated contribution of both morphemes. In the second, the lexicalized reduplicative already satisfies the reduplicative condition required by R_{A2}.

The idea that triplication involves two separate morphemes, rather than repeated applications of the same reduplication process, opens the door to a deeper morphological analysis. In particular, it raises the question of how these two morphemes interact with the base and with each other phonologically and morphologically. One key consequence of this hypothesis is that each morpheme may have its own set of allomorphs, with distinct phonological properties and selectional restrictions. This observation helps explain patterns of identity avoidance and tonal alternation that are otherwise difficult to capture. The next subsection builds on this hypothesis by formalizing the role of allomorphy in shaping the surface forms of triplication, particularly how R_{A2} and $R_{A1,\ B2,\ C1}$ select their forms depending on the base tone and structure.

5.8.2 Allomorphy selection

The pattern of identity avoidancce and tone harmony in triplication can be explained through the lens of allomorphy selection. Each of the two reduplicant morphemes, R_{A2} and $R_{A1,\ B2,\ C1}$, has its own allomorphs and selectional restrictions. These restrictions determine which combinations of base and reduplicant are permitted.

Specifically, I propose that the intensifying morpheme R_{A2} has two allomorphs: a prefixal allomorph that attaches to A1 bases, and a suffixal allomorph that occurs with all other bases. Prefix-suffix allomorphy in a single morpheme is not unusual, as seen in languages such as Nakanai (an Austronesian language in Papua New Guinea)'s nominalizer and Biak (an Austronesian language in Indonesia)'s third-person subject marking (Kalin, 2022). For each allomorph, there is a distinct ALIGN that they follow, presented below. In other words, morphology chooses the appropriate allomorph (prefix versus suffix), and once that choice is made, the phonological ALIGN constraints place it at the correct edge.

Similar to other reduplication processes, these alignment constraints will be dominant.

(38) a. R_{A2} prefix: ALIGN-L(R_{A2}, ω), ALIGN-R(Base, ω)

b. R_{A2} suffix: ALIGN-R(R_{A2}, ω), ALIGN-L(Base, ω)

For ineffable A2 bases, R_{A2} selects the base of the structure *base + reduplicant*. However, it would create an identity relationship where both the R_{A2} syllable and the base syllable have A2 tone. This identity relationship is systematically avoided in the language (see Section 3.5 for the avoidance pattern in triplication).

Consequently, each allomorph of R_{A2} leads to different placements of $R_{A1,\ B2,\ C1}$. For R_{A2} prefix, only R_{C1} associates with it, in which this $R_{A1,\ B2,\ C1}$ is also a prefix as a consequence. However, between R_{A2} and R_{C1}, R_{A2} is closer to the base than R_{C1} is. Thus, R_{C1} is the leftmost syllable in a trisyllabic reduplicative, and the base is the rightmost one; the ALIGN that governs the placement of R_{C1} (i.e., (39)) will dominate that of R_{A2}, as shown in (40) with an accompanied tableau.

(39) R_{C1} prefix: ALIGN-L(R_{C1}, ω), ALIGN-R(Base, ω)

(40) ALIGN-R(Base, ω), ALIGN-L(R_{C1}, ω) $\gg$ ALIGN-L(R_{A2}, ω)

BASE + R_{A2} + R_{C1}	BASE-R(ω)	R_{C1}-L(ω)	R_{A2}-L(ω)
a. ☞ C1.A2.BASE			*
b. BASE.A2.C1	*!*	*!*	*
c. A2.C1.BASE		*!	
d. A2.BASE.C1	*!	*!*	

Similar argument can be made for R_{A1} and R_{B2}, such that R_{A2} suffix leads to having R_{A1} and R_{B2} suffixes. Since R_{A2} is suffixal with regard to the base, R_{A1} and R_{B2} are the rightmost syllable in the word, while the base being the leftmost. The ALIGN that governs the placement of R_{A1} and R_{B2} (i.e., (41a)–(b)) will dominate that of R_{A2}, as shown in (42) with an accompanied tableau.[6]

(41) a. R_{A1} suffix: ALIGN-R(R_{A1}, ω), ALIGN-L(Base, ω)

b. R_{B2} suffix: ALIGN-R(R_{B2}, ω), ALIGN-L(Base, ω)

(42) ALIGN-L(Base, ω), ALIGN-R(R_{A1}, ω), ALIGN-R(R_{B2}, ω) $\gg$ ALIGN-L(R_{A2}, ω)

BASE + R_{A2} + R_{A1}	R_{A1}-R(ω)	BASE-L(ω)	R_{A2}-R(ω)
a. ☞ BASE.A2.A1			*
b. A1.A2.BASE	*!*	*!*	*
c. BASE.A1.A2	*!		
d. A1.BASE.A2	*!*	*!	

Regarding $R_{A1,\ B2,\ C1}$, it alternates between A1, B2, and C1. For these varying tones, the tone specification in the reduplicant can fall into three levels: (1) no specification (phonolog-

[6] R_{A1} is used in the tableau, but the alignment of R_{B2} should also be the same.

ically derived), (2) partial specification (e.g., only contour or register is fixed), and (3) full specification (both tone and segments fixed, forming an allomorph).

5.8.2.1 A1

When there is no specification, the tone is fully derived in the phonology. This derivation can follow two possible rankings:

(43) MAX-IO ≫ *L/OCP (T), NOCONTOUR ≫ MAX-BR

Both ranking use NOCONTOUR for the flat tone to emerge. The first ranking uses *L to penalize low tones (i.e., −H tone) and allows a high tone, in this case A1, to emerge (Bickmore, 2000). Another way to approach this is to have OCP (T) instead of *L. This constraint avoids adjacent identical tones, ruling out A2 due to identity with the tone R_{A2}. OCP (T) is based on the OCP constraint proposed by Golston (1996), which militates against adjacent identical elements. This constraint, in turn, is derived from the *Obligatory Contour Principle* (OCP) introduced in Leben (1973), which prohibits adjacent identical tones.

When there is partial specification of the contour (denoted Aα), the ranking still involves either *L or OCP (T). The NOCONTOUR constraint is not needed because the flat contour is already specified in the input.

(44) MAX-IO ≫ *L/OCP (T) ≫ MAX-BR

When the tone is partially specified for register only (e.g., T1 for +H), there is only one approach, using NOCONTOUR ensures a flat tone surfaces. There is no need for *L nor OCP because if the reduplicant is already specified with a H register, it inherently satisfies *L and OCP.

(45) MAX-IO ≫ NOCONTOUR ≫ MAX-BR

When fully specified as A1, MAX-IO simply dominates MAX-BR to enforce complete faithfulness to the input, making the tone part of the reduplicant's lexical representation. Thus, the result is a constraint ranking similar to that of full reduplication in the attenuation process (i.e., (16))

(46) Phono, Fill, Align-R(Ft, H(Ft)), ParseSyll, FtBin, Max-IO ≫ Max-BR

Among these three patterns — no specification, partial specification, and full specification — the full specification approach offers the most theoretically consistent analysis. While both the no specification and partial specification approaches require additional markedness constraints like *L, Ocp (T), or NoContour to derive the correct tone patterns, the full specification approach relies simply on Max-IO dominating Max-BR (46). Crucially, this ranking parallels the constraint rankings already established for attenuation, pejorative, intensification, and AABB reduplication within the language. Adopting the full specification analysis therefore maintains a unified constraint ranking across all reduplicative processes in the same prosodic domain. In contrast, the other approaches would introduce ranking paradoxes where certain markedness constraints must sometimes be dominated (in full reduplication) and sometimes dominate (when phonologically deriving the reduplicant's tone). By treating the A1 tone as lexically specified in the reduplicant, we eliminate these conflicts and provide a more coherent overall analysis of Vietnamese prosodic morphology.

5.8.2.2 B2

For B2, a similar set of derivations can be proposed. With no specification, there are two possible rankings account for the surface form.

(47) Max-IO ≫ Ocp (T)/*H ≫ Max-BR

In the first ranking, adjacent identical tones are avoided. However, for Ocp (T) to work here, we have to change its scope from identical tones in adjacent syllables in a prosodic word to those in any *pairing of syllables* in the prosodic word. For instance, if Ocp works for first syllable-second syllable and second syllable-third syllable in a linear order, it should now be applicable to first syllable-third syllable as well. Note that the change of scope in Ocp (T) does not affect the ranking for A1 in (43) because the base and the reduplicant in question never have a matching tone.

A second account appeals to the markedness constraint *H, which penalizes high tones

(Bickmore, 2000). The base has B1, and with MAX-IO ≫ *H, it remains unaltered. In the case of B2, its base already has B1, meaning *H will always be violated at most once; if the reduplicant has B1 instead of B2, it will have an additional violation. Although A2 is a possible outcome under this ranking, tone-coda co-occurrence constraints make it less optimal, as A2 would require a homorganic nasal coda in the reduplicant, further violating MAX-BR.

In partial specification cases with only a contour (Bα), register assignment follows the same logic as in the no-specification scenario and hence has the same constraint ranking. When only register is specified (−H), surface forms like A2, B2, and C2 are possible. However, phonotactic constraints again rule out A2 and C2 due to required coda alternations. Since B2 can surface without phonotactic violations, it incurs fewer MAX-BR violations, making it the optimal form. Fully specified B2 reduplicants are protected by the high-ranking MAX-IO constraint. This means that both latter specifications have the same constraint ranking as that of full reduplication in attenuation and AABB.

(48) PHONO, FILL, ALIGN-R(Ft, H(Ft)), PARSESYLL, FTBIN, MAX-IO ≫ MAX-BR

As with the A1 reduplicant analysis, the full specification approach for B2 offers considerable theoretical advantages. While the no-specification approach requires either extending the scope of OCP (T) to non-adjacent syllables or introducing *H constraint, and the partial specification approach involves complexities with register assignment and phonotactic interactions, the full specification approach simply relies on the established ranking MAX-IO ≫ MAX-BR (48). This ranking aligns with the constraint rankings already established for attenuation and AABB reduplication processes. The consistency of this pattern across different reduplicative processes suggests that treating B2 as fully specified in the reduplicant's lexical representation may be the most parsimonious approach.

5.8.2.3 C1

For C1, both full-specified and contour-specified reduplicants share the same constraint ranking: MAX-IO ≫ MAX-BR. In full specification, the argument mirrors that of A1 and

B2. In the contour-specified case, C2 is a competitor, but it deletes not only the base's contour but also the base's register in the reduplicant, violating MAX-BR twice. Hence, C1 is preferred.

When tones are unspecified or only register-specified, the following ranking accounts for the outcome.

(49) MAX-IO ≫ OCP (T), IDENT-BR ([αCG]) ≫ MAX-BR

OCP (T) eliminates repeated tones, leaving only A1 and A2 as possibilities. B2 and C2 are excluded because they mismatch the base in both register and contour. To further ensure that C1 is chosen over B1 (which shares the same contour), we can introduce the ad-hoc constraint IDENT-BR ([αCG]), which favors base-reduplicant pairs that agree in the α value of the [Constricted Glottis] feature. Since A1 and C1 lack [CG], i.e., [−CG], while B1 possesses it (Brunelle, 2009), this constraint favors C1.

The ranking works as follows, as exemplified by the input /nɤwA1/ 'brown' with an unspecified reduplicant RED. The optimal output is candidate (a) with C1 in the reduplicant, which deletes the base's tone contour in the reduplicant. Candidate (b) also violates MAX-BR as much as the optimal output does, is eliminated since it violates other markedness constraints that dominate MAX-BR. Candidate (c) with the identical base and reduplicant is eliminated because they have a pair of identical tones in the output. Candidate (d) is eliminated for violating the dominant MAX-IO.

(50) Input: /nɤwA1/ 'brown'

Red + R_{A2} + nɤwA1	Max-IO	Ocp (T)	Ident-BR ([αCg])	Max-BR
a. ☞ nɤwC1.nɤwA2.nɤwA1				*
b. nɤwB1.nɤwA2.nɤwA1			*!	*
c. nɤwA1.nɤwA2.nɤwA1		*!		
d. nɤwC1.nɤwA2.nɤwC1	*!			

The preceding analyses of the three tonal allomorphs (R_{A1}, R_{B2}, and R_{C1}) show that the selectional restrictions on $R_{A1,\ B2,\ C1}$ demonstrate that morphology consults phonology, i.e., the tone of the base, when the allomorph chooses the appropriate base. The high-ranking Max-IO and Align constraints show that phonology cannot freely alter a lexically specified allomorph once selected.

The relationship between morphology and phonology reveals a noticeable ranking paradox, which lies in Ocp. Ocp operates across the tone alternations in triplication, and there is no separability effect in triplication, meaning the trisyllabic form behaves as a single prosodic word (PrWd). This means that Ocp is operating in the domain of PrWd. Compare to AABB and the AB base in that process (see Section 3.2), there is separability effect which indicates that A and B in AB and AA, BB in AABB form are all separate PrWds. This raises a paradox: if Ocp is active within PrWds in triplication, it should also be active in AABB reduplication, where the domain is similarly defined. However, identical adjacent syllables are allowed in AABB, contradicting Ocp's supposed influence. This paradox implies a ranking conflict, which is Ocp dominates Max-BR in triplication but is dominated in AABB. These ranking contradictions highlight the inadequacy of underspecified representations for reduplicants.

In addition to resolving internal ranking issues, the analysis of these three reduplicant patterns reveals a significant advantage of the full specification approach over alternatives. Both no-specification and partial-specification analyses require introducing markedness con-

straints that would create ranking paradoxes when considered alongside other reduplication processes in Vietnamese. These constraints would need to dominate MAX-BR in triplication to derive the correct tonal patterns, yet must be dominated by MAX-BR in processes like full reduplication. The full specification approach avoids these ranking conflicts entirely by treating each reduplicant tone (A1, B2, and C2) as lexically specified allomorphs of the $R_{A1,\ B2,\ C1}$ morpheme. This approach maintains a uniform constraint ranking across all reduplicative processes within the same prosodic domain: MAX-IO consistently dominates MAX-BR. Thus, understanding triplication as involving tone-specific allomorphs rather than phonological derivation provides the most coherent analysis of Vietnamese tonal reduplication as a whole.

This allomorphy-based analysis not only avoids ranking paradoxes but also aligns with typological observations: true triplication is rare cross-linguistically and often reducible to multiple affixation or allomorph selection. Vietnamese triplication, then, exemplifies a morphological strategy encoded in the lexicon rather than a product of recursive reduplication. The desired ranking for each allomorph is presented below. One thing to note is that each allomorph is subject to their own set of alignment constraints; phonological-material wise, it is lexically specified.

(51) A1:

PHONO, FILL, ALIGN-R(Ft, H(Ft)), PARSESYLL, FTBIN, MAX-IO, ALIGN-L(Base, ω), ALIGN-R(R_{A1}, ω), ALIGN-R(R_{B2}, ω), CONTIGBASE, CONTIGRED, $\gg$ ALIGN-L(R_{A2}, ω), MAX-BR

(52) B2:

PHONO, FILL, ALIGN-R(Ft, H(Ft)), PARSESYLL, FTBIN, MAX-IO, ALIGN-L(Base, ω), ALIGN-R(R_{A1}, ω), ALIGN-R(R_{B2}, ω), CONTIGBASE, CONTIGRED $\gg$ ALIGN-L(R_{A2}, ω), MAX-BR

(53) C1:

PHONO, FILL, ALIGN-R(Ft, H(Ft)), PARSESYLL, FTBIN, MAX-IO, ALIGN-R(Base, ω), ALIGN-L(R_{C1}, ω), CONTIGBASE, CONTIGRED $\gg$ ALIGN-L(R_{A2}, ω), MAX-BR

Triplication may be better understood not as a base repeating twice but as two reduplicative allomorphs attached to a single base. If it was base being copied n times, the form of the reduplicant is phonologically derived. However, under an allomorphy-based analysis, the number of possible reduplicants is inherently limited by the inventory of available allomorphs in the lexicon, each associated with a specific morphological or semantic function. This allomorphy-based approach accounts for tone alternations, identity avoidance, and the distribution of reduplicants in triplication, tying together the phonological and morphological aspects of Vietnamese reduplication.

5.9 Interim summary

This chapter has provided an OT analysis of several reduplication patterns in Vietnamese, each serving distinct functions: attenuation, intensification of polysyllabic adjectives ([-aː] reduplication), exclamation (AABB), and intensification of monosyllabic adjectives (triplication). While each pattern could be analyzed in isolation and as a phonological derivation through constraint interactions, doing so risks positing multiple independent constraint rankings within a single language, which is theoretically undesirable. Processes involving full reduplication, like AABB, should be analyzed in tandem with their tone patterns, as tone provides critical evidence for how these forms are derived.

I suggest a unified ranking across all reduplication types by adopting an allomorphy approach. Each reduplicant is treated as having a set of lexically specified allomorphs that select the base or base-reduplicant combination they attach to (Fabb, 1988) and hence is not phonologically derived. This perspective helps avoid complications that arise under a reduplicative analysis, especially those involving tone identity or alternation that cannot be easily explained if we assume multiple rounds of phonological copying.

While I draw on tools from OT to formalize the constraints and interactions that govern

these patterns, my argument is not theory-internal to OT. Instead, OT serves here as a framework to express broader generalizations about morphophonological behavior, particularly the role of tone in morphophonological processes like reduplication and allomorphy.

CHAPTER 6

CONCLUSION

This thesis has developed a unified, OT-based analysis of Vietnamese reduplication, focusing on previously under-analyzed data from Southern Vietnamese. The analysis offers an account of various reduplicative patterns — from well-documented disyllabic forms to complex cases such as triplication and AABB — while addressing tone behavior and avoidance effects across processes.

Three main contributions have emerged from this research. First, this thesis provides substantial new empirical data on previously under-documented reduplication patterns, notably triplication and AABB. This novel data, gathered through rigorous fieldwork involving native speakers, fills an important gap in Vietnamese linguistic documentation and provides the empirical foundation necessary for robust theoretical analysis.

Second, this thesis proposes a formalization of Vietnamese reduplication under OT, arguing that the so-called triplication in Vietnamese emerges not from repeated reduplication but from attaching *two distinct reduplicative morphemes*, each possessing its own set of allomorphs. In other words, Vietnamese triplication is morphological, such that (1) morphology chooses the allomorph, which is triggered by phonological properties of the base (especially tone), and (2) phonology merely enforces the choice through high-ranking Align and Max-IO keep the selected allomorph at its lexically stored edge and preserve its tone. As a whole, Vietnamese reduplication is best understood through the selection of *lexically specified reduplicant morphemes* rather than phonological derivation. This perspective offers a unified explanation for tone harmony, avoidance effects, and prosodic alignment across various reduplication processes.

Third, avoidance patterns are offered a description, and they can be understood as a consequence of the allomorph selecting a particular base to attach to, which may resemble identity avoidance on the surface. The analysis utilizes the OT framework to formalize the argument; however, it is not dependent on the theory.

This analysis opens several directions for future research. Particularly, a more robust morphophonological account may be needed. In verb attenuation, the lack of observed allomorphy, in contrast to the adjectival case, suggests that distinct morphemes may underlie the reduplication of verbs and adjectives. The use of a placeholder constraint PHONO to simplify Vietnamese phonotactics invites a more detailed exploration of tone–coda co-occurrence restrictions. The preference for B1 reduplicants in pejorative constructions, even with non-B1 bases, may reflect a tonal markedness pattern and warrants empirical testing. The separability effect and internal movement of reduplicants in AABB reduplication suggest the possibility of two distinct reduplication processes at work. If this is the case, it could help explain both the separability effect observed in AABB reduplicatives and the apparent internal movement of reduplicants. Importantly, this movement cannot simply be attributed to compound structure, as compounds block internal movement, ruling out any analysis in which a portion of the compound (i.e., the reduplicant) is moved. Finally, the separability seen in AABB, pejorative, and intensification constructions remains insufficiently addressed, especially with respect to stress placement and prosodic structure, and presents an opportunity for more detailed analysis.

BIBLIOGRAPHY

Agbayani, B. (1997). *Phonological Constraints and Base-Reduplicant Non-Identity: The Case of Vietnamese* [Master's thesis, University of California, Irvine].

Alderete, J., Beckman, J., Benua, L., Gnanadesikan, A., McCarthy, J., & Urbanczyk, S. (1999). Reduplication with Fixed Segmentism. *Linguistic Inquiry*, *30*(3), 327–364. https://doi.org/10.1162/002438999554101

An, Y.-r. (2012). The gradient OCP: Evidence from Korean reduplication. In Rebeka Campos-Astorkiza & Jon Franco (Eds.), *Papers in linguistics by the BIDE generation* (pp. 33–53, Vol. 46). UPV/EHU Press.

Anttila, A. (2002). Morphologically conditioned phonological alternations. *Natural Language and Linguistic Theory*, *20*(1), 1–42. https://doi.org/10.1023/A:1014245408622

Arto Anttila. (1997). Deriving Variation from Grammar. In Frans L. Hinskens, Roeland van Hout, & W. Leo Wetzels (Eds.), *Current Issues in Linguistic Theory* (pp. 35–68, Vol. 146). John Benjamins Publishing Company. https://roa.rutgers.edu/files/63-0000/63-0000-ANTTILA-0-0.PDF

Bickmore, L. S. (2000). Downstep and fusion in Namwanga. *Phonology*, *17*(3), 297–331. https://doi.org/10.1017/S0952675701003943

Brunelle, M. (2009). Tone perception in Northern and Southern Vietnamese. *Journal of Phonetics*, *37*(1), 79–96. https://doi.org/10.1016/j.wocn.2008.09.003

Brunelle, M. (2017). Stress and phrasal prominence in tone languages: The case of Southern Vietnamese [Edition: 2017/02/16 Publisher: Cambridge University Press]. *Journal of the International Phonetic Association*, *47*(3), 283–320. https://doi.org/10.1017/S0025100316000402

Cao, X.-H. (2003). *Tiếng Việt: Mấy vấn đề ngữ âm, ngữ pháp, ngữ nghĩa (Vietnamese: A few phonetic, grammatical and semantic problems*. The Ministry of Education; Training.

Cao, X.-H. (2007). *Ngữ âm tiếng Việt (Vietnamese Phonetics)*. VNU.

Carleton, T., & Myers, S. (1994). *On tonal transfer*. https://roa.rutgers.edu/files/16-0594/16-0594-CARLETON-0-0.PDF

Chen, M. Y. (2000). *Tone sandhi: Patterns across Chinese dialects* (Vol. 92). Cambridge University Press.

Clements, G. N., & Hume, E. V. (1995). The internal organization of speech sounds. In J. A. Goldsmith (Ed.), *The Handbook of Phonological Theory* (pp. 245–306). Blackwell.

Corbin, D., & Plénat, M. (1992). NOTE SUR L'HAPLOLOGIE DES MOTS CONSTRUITS [Publisher: Armand Colin]. *Langue Française*, (96), 101–112. http://www.jstor.org/stable/41558606

Cu, D.-T., Hoang, V.-T., & Nguyen, N.-T. (1977). *Ngữ âm học tiếng Việt hiện đại (Phonology of the modern Vietnamese)*. The Ministry of Education; Training.

de Lacy, P. (1999). Morphological haplology and correspondence. In *University of Massachusetts Occasional Papers: Papers from the 25th Reunion* (pp. 51–88, Vol. 24). GLSA Publications. https://roa.rutgers.edu/files/298-0299/roa-298-lacy-1.pdf

Diep, Q.-B., & Hoang, V.-T. (1996). *Ngữ pháp tiếng Việt (Vietnamese Grammar)* (Vol. 1). The Ministry of Education; Training.

Fabb, N. (1988). English Suffixation Is Constrained Only by Selectional Restrictions [Publisher: Springer]. *Natural Language & Linguistic Theory*, *6*(4), 527–539. http://www.jstor.org/stable/4047592

Fitzgerald, C. M. (2000). Vowel Hiatus and Faithfulness In Tohono O'Odham Reduplication. *Linguistic Inquiry*, *31*(4), 713–722. https://doi.org/10.1162/002438900554532

Golston, C. (1996). Direct Optimality Theory: Representation as Pure Markedness. *Language*, *72*(4), 713–748. https://doi.org/10.2307/416100

Hayes, B. (1989). The prosodic hierarchy in meter. In P. Kiparsky & G. Youmans (Eds.), *Phonetics and Phonology* (pp. 201–260, Vol. 1). Academic Press.

Hoang, V.-H. (1985). *Từ láy trong Tiếng Việt (Vietnamese Reduplication)*. Social Sciences Publishing House.

Hoang, V.-H. (1998). *Từ điển từ láy tiếng Việt (Vietnamese Reduplication Dictionary)*. Social Sciences Publishing House.

Hoang, V.-H., Ha, Q.-N., & Nguyen, V.-K. (2008). *Từ tiếng Việt (Vietnamese word)*. Social Sciences Publishing House.

Hsiao, Y. E. (2018). An Analysis Of Taiwan Southern Min Tetrasyllabic Adjective Reduplication. *Journal of Chinese Linguistics*, *46*(1), 123–148. https://doi.org/10.1353/jcl.2018.0004

Hyman, L. M. (2014). Coda constraints on tone. In S. Bendjaballah, N. Faust, M. Lahrouchi, & N. Lampitelli (Eds.), *Language Faculty and Beyond* (pp. 103–122, Vol. 12). John Benjamins Publishing Company. https://doi.org/10.1075/lfab.12.09hym

Inkelas, S., Orgun, C. O., & Zoll, C. (1997). The Implications of Lexical Exceptions for

the Nature of Grammar. In I. Roca (Ed.), *Derivations and Constraints in Phonology* (pp. 393–418). Clarendon Press. https://doi.org/10.1093/oso/9780198236894.003.0013

Inkelas, S., & Zoll, C. (2005). *Reduplication: Doubling in Morphology* (1st ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511627712

Inkelas, S., & Zoll, C. (2007). Is grammar dependence real? A comparison between cophonological and indexed constraint approaches to morphologically conditioned phonology. *Linguistics*, *45*(1). https://doi.org/10.1515/LING.2007.004

Kalin, L. (2022). Infixes really are (underlyingly) prefixes/suffixes: Evidence from allomorphy on the fine timing of infixation. *Language*, *98*(4), 641–682. https://doi.org/10.1353/lan.2022.0017

Kennard, C. H. (2004). Copy but don't repeat: The conflict of dissimilation and reduplication in the Tawala durative. *Phonology*, *21*(3), 303–323. https://doi.org/10.1017/S0952675704000296

Kennedy, R. (2008). Evidence for Morphoprosodic Alignment in Reduplication. *Linguistic Inquiry*, *39*(4), 589–614. https://doi.org/10.1162/ling.2008.39.4.589

Kurisu, k. (2001). The Phonology of Morpheme Realization. https://doi.org/10.7282/T3R49PKM

Leben, W. (1973). *Suprasegmental Phonology* [PhD Thesis]. MIT. http://www.ai.mit.edu/projects/dm/theses/leben73.pdf

Lin, H.-s. (2015). Chengdu Reduplication: An Optimality Theoretic Analysis. *Language and Linguistics*, *16*(6), 843–872. https://doi.org/10.1177/1606822X15602613

Lombardi, L. (1997). Coronal sonorants and markedness. In *Maryland Working Papers in Linguistics* (pp. 156–175, Vol. 5). University of Maryland.

McCarthy, J. J., & Prince, A. (1994). The emergence of the unmarked: Optimality in prosodic morphology. In *Proceedings of the North East Linguistics Society 24* (Vol. 18). GLSA, UMASS. https://scholarworks.umass.edu/linguist_faculty_pubs/18

McCarthy, J. J., & Prince, A. (1995). Faithfulness and reduplicative identity. In J. Beckman, L. W. Dickey, & S. Urbanczyk (Eds.), *Papers in Optimality Theory* (pp. 249–384, Vol. 10). GLSA, UMASS. https://scholarworks.umass.edu/linguist_faculty_pubs/10

Michaud, A. (2005). Final consonants and glottalization: New perspectives from Hanoi Vietnamese. *Phonetica*, *61*(2-3), 119–146. https://doi.org/10.1159/000082560

Ngo, T.-N. (1984). *The syllabeme and patterns of word formation in Vietnamese*. New York University.

Nguyen, D.-H. (1997). *Vietnamese*. John Benjamins.

Nguyen, P.-P. (1998). Các phương diện hình thái học tiếng Việt (Aspects of Vietnamese Morphology). *Hội thảo Quốc tế Việt Nam học (Internal Conference on Vietnamology)*, *1*, 139–150. https://repository.vnu.edu.vn/bitstream/VNU_123/20295/1/018.pdf

Nguyen, T., & Ingram, J. (2007). Acoustic and perceptual cues for compound-phrasal contrasts in Vietnamese. *The Journal of the Acoustical Society of America*, *122*. https://doi.org/10.1121/1.2747169

Nguyen, T., & Ingram, J. C. L. (2006). Reduplication and word stress in Vietnamese.

Nguyen, V.-L., & Edmondson, J. A. (1997). Thanh điệu và chất giọng trong tiếng Việt hiện đại (phương ngữ Bắc Bộ): Khảo sát thực nghiệm (The tones and voice quality in modern northern Vietnamese: Instrumental case studies). *Ngôn ngữ (Language)*, *1*, 1–16.

Orgun, C. O. (1996). *Sign-Based Morphology and Phonology: With special attention to Optimality Theory* [PhD Thesis]. University of California, Berkeley. https://roa.rutgers.edu/files/171-0197/171-0197-ORGUN-0-0.PDF

Parker, S. (2018). Central vs. Back Vowels. *Work Papers of the Summer Institute of Linguistics, University of North Dakota Session*, *44*(1). https://doi.org/10.31356/silwp.vol44.05

Pham, A. H. (2004). *Vietnamese Tone: A New Analysis* (L. Horn, Ed.; 1st ed.). Routledge. https://doi.org/10.4324/9780203500088

Pham, A.-H. (2001). Vietnamese tone: Tone is not pitch [Section: University of Toronto Linguistics PhD students' Dissertations]. *Toronto Working Papers in Linguistics*. Retrieved October 28, 2024, from https://twpl.library.utoronto.ca/index.php/twpl/article/view/6507

Pham, A.-H. (2009). The identity of non-identified sounds: Glottal stop, prevocalic /w/ and triphthongs in Vietnamese [Section: Articles]. *Toronto Working Papers in Linguistics*, *34*. Retrieved October 28, 2024, from https://twpl.library.utoronto.ca/index.php/twpl/article/view/6803

Pham, A.-H., & Pham, A.-A. (2020). Productive Reduplication in Southern Vietnamese [i]. *Journal of the Southeast Asian Linguistics Society*, *13*(1). Retrieved October 28, 2024, from https://link.gale.com/apps/doc/A651817246/AONE?u=anon~e7c7873d&sid=googleScholar&xid=afce68dd

Prince, A., & Smolensky, P. (2002). Optimality Theory: Constraint Interaction in Generative Grammar. https://doi.org/https://doi.org/10.7282/T34M92MV

Pulleybank, D. (1997). Optimality Theory and Features. In *Optimality theory: An overview* (pp. 59–101, Vol. 59). Blackwell. https://www.researchgate.net/profile/Douglas-Pulleyblank/publication/247074150_Optimality_theory_and_features/links/540d4d590cf2f2b29a383763/Optimality-theory-and-features.pdf

Russell, K. (1995). *Morphemes and candidates in Optimality Theor*. https://roa.rutgers.edu/files/44-0195/44-0195-RUSSELL-0-0.PDF

Selkirk, E. O. (1978). On prosodic structure and its relation to syntactic structure. In T. Fretheim (Ed.), *Nordic Prosody II* (pp. 111–140). Tapir.

Srichampa, S. (2002). Vietnamese verbal reduplication. In *Collected papers on Southeast Asian and Pacific languages* (pp. 37–47). Pacific Linguistics. https://openresearch-repository.anu.edu.au/server/api/core/bitstreams/ebd2dcac-aee0-491d-971f-1e3bec97cc1a/content

Thompson, L. C., & O'Harrow, S. D. (1988). *A Vietnamese Reference Grammar*. University of Hawaii Press. https://books.google.com/books?id=XN7SHNeZ_ksC

Tsay, J., & Myers, J. (1996). Taiwanese Tone Sandhi as Allomorph Selection. *Annual Meeting of the Berkeley Linguistics Society*, *22*(1), 395. https://doi.org/10.3765/bls.v22i1.1347

Vo, X.-H. (1997). *Nghiên cứu về chức năng của thanh điệu tiếng Việt (theo phương pháp định lượng) (Quantitative Studies on Vietnamese Tones)* [PhD Thesis]. Hanoi University of Education.

Vu, S. X. (2007). A unified analysis of some Vietnamese reduplication forms. In M. Alves, P. Sidwell, & D. Gil (Eds.), *SEALS VIII Papers from the 8th Annual Meeting of the Southeast A* (pp. 165–191). Pacific Linguistics, Research School of Pacific; Asian Studie. http://sealang.net/sala/archives/pdf8/vu1998unified.pdf#page=16.00

Wivell, G., Miatto, V., Karakaş, A., Kostyszyn, K., & Repetti, L. (2024). All about ablaut: A typology of ablaut reduplicative structures. *Linguistic Typology*, *28*(3), 505–536. https://doi.org/10.1515/lingty-2023-0018

Wolf, M. (2005). For an Autosegmental Theory of Mutation. *Proceedings of the West Coast Conference on Formal Linguistics*. https://doi.org/10.7282/T3CZ3B3S

Yip, M. (1995). Identity Avoidance in Phonology and Morphology. *Proceedings of the Conference on Morphology and its relation to Syntax and Phonology*. https://roa.rutgers.edu/files/82-0000/roa-82-yip-3.pdf#page=20.73

Zhang, J., & Lai, Y. (2007). Two aspects of productivity in Taiwanese Double Reduplication. *Kansas Working Papers in Linguistics*. https://doi.org/10.17161/KWPL.1808.1786

APPENDIX A

GENERAL STRUCTURE OF REDUPLICATION

Legend

X	Base	red	Verb
$\overline{\text{X}}$	RED	blue	Adjective
N	Nasal	purple	Verb + Adjective
T	Tone contour	black	Noun + Verb + Adjective
α	Tone register		Novel words
green	Noun		

Type	**Base**	**Base + Reduplicant**
Attenuation	X	$\overline{\text{X}}$.X
	X	$\overline{\text{X}}$.X
		$\overline{\text{X}}_{\text{A}\alpha}$.X ($\alpha$ matches that of X)
Pejorative	X	X.$\overline{\text{X}}_{\text{iəkB1/B2}}$
	X.Y	X.Y.$\overline{\text{X}}$.$\overline{\text{Y}}_{\text{iəkB1/B2}}$
	X-T1	X.$\overline{\text{X}}_{\text{iəkB}\alpha}$
	X.Y T1	X.Y.$\overline{\text{X}}$.$\overline{\text{Y}}_{\text{iəkB}\alpha}$
	X(.Y)$_{\text{iək}}$	$\varnothing$
	X	X.$\overline{\text{X}}_{\text{iəkB}\alpha}$ (less alternation)
	X.Y	X.Y.$\overline{\text{X}}$.$\overline{\text{Y}}_{\text{iəkB}\alpha}$ (less alternation)
	X-T1	X.$\overline{\text{X}}_{\text{iəkB}\alpha}$
	X(.Y)$_{\text{iək}}$	$\varnothing$

Exclamation	X.Y	X.$\overline{\text{X}}$.$\overline{\text{Y}}$.Y
		$\overline{\text{X}}$.X.Y.$\overline{\text{Y}}$
		X.$\overline{\text{X}}$.Y.$\overline{\text{Y}}$
		$\overline{\text{X}}$.X.$\overline{\text{Y}}$.Y
	X.X	∅
	X.Y	$\overline{\text{X}}$.X.Y.$\overline{\text{Y}}$
		X.$\overline{\text{X}}$.Y.$\overline{\text{Y}}$
		$\overline{\text{X}}$.X.$\overline{\text{Y}}$.Y
		X.$\overline{\text{X}}$.$\overline{\text{Y}}$.Y
	$\text{X.X}_{\text{ɛC}\alpha}$	$\overline{\text{X}}$.$\overline{\text{X}}_{\text{a:A}\alpha}$.$\text{X.X}_{\text{ɛC}\alpha}$
	$\text{X.X}_{\text{a:ŋA}\alpha}$	$\overline{\text{X}}$.X.Y.$\overline{\text{Y}}$
		X.$\overline{\text{X}}$.Y.$\overline{\text{Y}}$
		$\overline{\text{X}}$.X.$\overline{\text{Y}}$.Y
		X.$\overline{\text{X}}$.$\overline{\text{Y}}$.Y
		$\overline{\text{X}}$.$\overline{\text{X}}_{\text{a:A}\alpha}$.$\text{X.X}_{\text{a:ŋA}\alpha}$
	$\text{X.X}_{\text{a:ŋB}\alpha}$	$\overline{\text{X}}$.$\overline{\text{X}}_{\text{a:A}\alpha}$.$\text{X.X}_{\text{a:ŋB}\alpha}$
	$\text{X.X}_{\text{a:pB2}}$	$\overline{\text{X}}$.X.Y.$\overline{\text{Y}}$
		X.$\overline{\text{X}}$.Y.$\overline{\text{Y}}$
		$\overline{\text{X}}$.X.$\overline{\text{Y}}$.Y
		X.$\overline{\text{X}}$.$\overline{\text{Y}}$.Y
		$\overline{\text{X}}$.$\overline{\text{X}}_{\text{a:A2}}$.$\text{X.X}_{\text{a:pB2}}$
	$\text{X.X}_{\text{iəkB}\alpha}$	$\overline{\text{X}}$.X.Y.$\overline{\text{Y}}$
		X.$\overline{\text{X}}$.Y.$\overline{\text{Y}}$
		$\overline{\text{X}}$.X.$\overline{\text{Y}}$.Y
		X.$\overline{\text{X}}$.$\overline{\text{Y}}$.Y

	X.X	∅
Intensification (1)	$X.Y_{T\alpha}$	$\overline{X}.\overline{Y}_{a\text{ː}A\alpha}.X.Y_{T\alpha}$
	$X.X_{i\text{ə}kB1}$	$X.\overline{X}$-aːA1.X.$X_{i\text{ə}kB1}$
	$X.Y_{T\alpha}$	$\overline{X}.\overline{Y}_{a\text{ː}A\alpha}.X.Y_{T\alpha}$
	$X.X_{i\text{ə}kB1}$	$\overline{X}.\overline{X}$-aːA1.X.$X_{i\text{ə}kB1}$
	$X.X_{a\text{ː}\eta A\alpha}$	X.$\overline{X}$.$\overline{Y}$.Y
		$\overline{X}$.X.Y.$\overline{Y}$
		X.$\overline{X}$.Y.$\overline{Y}$
		$\overline{X}$.X.$\overline{Y}$.Y
		$\overline{X}.\overline{X}_{a\text{ː}A\alpha}$.X.$X_{a\text{ː}\eta A\alpha}$
	$X.X_{a\eta A\alpha}$	X.$\overline{X}$.$\overline{Y}$.Y
		$\overline{X}$.X.Y.$\overline{Y}$
		X.$\overline{X}$.Y.$\overline{Y}$
		$\overline{X}$.X.$\overline{Y}$.Y
		$\overline{X}.\overline{X}_{a\text{ː}A\alpha}$.X.$X_{a\eta A\alpha}$
	$X.X_{a\text{ː}pB\alpha}$	X.$\overline{X}$.$\overline{Y}$.Y
		$\overline{X}$.X.Y.$\overline{Y}$
		X.$\overline{X}$.Y.$\overline{Y}$
		$\overline{X}$.X.$\overline{Y}$.Y
		$\overline{X}.\overline{X}_{a\text{ː}A\alpha}$.X.$X_{a\text{ː}pB\alpha}$
	$X.X_{\varepsilon\text{ː}C\alpha}$	X.$\overline{X}$.$\overline{Y}$.Y
		$\overline{X}$.X.Y.$\overline{Y}$
		X.$\overline{X}$.Y.$\overline{Y}$
		$\overline{X}$.X.$\overline{Y}$.Y
		$\overline{X}.\overline{X}_{a\text{ː}A\alpha}$.X.X.$X_{\varepsilon\text{ː}C\alpha}$

Intensification (2)/ Triplication	X	$X.\overline{X}_{A2}.\overline{X}_{A1}$
	X_{B1}	$X_{B1}.\overline{X}_{A2}.\overline{X}_{B2}$ (X has an obstruent coda)
	X.Y	$X.\overline{Y}_{A\alpha}.Y$
	X	$X.\overline{X}_{A2}.\overline{X}_{A1}$
	X_{B1}	$X_{B1}.\overline{X}_{A2}.\overline{X}_{B2}$ (X has an obstruent coda)

APPENDIX B

ADDITIONAL DATA

B.1 Attenuation

Verb

(1)

a.	[ŋiA1]	'to suspect'	[ŋiA1.ŋiA1]	'suspect (ATT)'
			*[ŋiA2.ŋiA1]	
b.	[ɲɨnA2]	'to look'	[ɲɨnA2.ɲɨnA2]	'look (ATT)'
			*[ɲɨnA1.ɲɨnA2]	
c.	[tʰɨtB1]	'to like'	[tʰɨtB1.tʰɨtB1]	'like (ATT)'
			*[tʰɨnA1.tʰɨtB1]	
d.	[liəkB1]	'to glance'	[liəkB1.liəkB1]	'glance (ATT)'
			*[liəŋA1.liəkB1]	

Adjective

(2) **A1 words**:

a.	[kaːwA1]	'tall'	[kaːwA1.kaːwA1]	'slightly tall'
			*[kaːwA2.kaːwA1]	
b.	[ɗɛŋA1]	'black'	[ɗɛŋA1.ɗɛŋA1]	'slightly black'
			*[ɗɛŋA2.ɗɛŋA1]	

(3) **A2 words**:

a.	[vaːŋA2]	'yellow'	[vaːŋA2.vaːŋA2]	'slightly yellow'
			*[vaːŋA1.vaːŋA2]	
b.	[jaːA2]	'old'	[jaːA2.jaːA2]	'slightly old'
			*[jaːA1.jaːA2]	

(4) **B1 words**:

a.	[tɨmB1]	'purple'	[tɨmB1.tɨmB1]	'slightly purple'
			[tɨmA1.tɨmB1]	
b.	[sopB1]	'porous; fluffy'	[sopB1.sopB1]	'slightly fluffy'
			[somA1.sopB1]	
c.	[saːkB1]	'close to'	[saːkB1.saːkB1]	'slightly close to'
			[saːŋA1.saːkB1]	
			*[saːkB1.saːŋA1]	(R is the rightmost σ)
			*[saː*kA1*.saːkB1]	(tone-coda co-occurrence)

(5) **B2 words**:

a.	[ɗɛpB2]	'pretty'	[ɗɛpB2.ɗɛpB2]	'slightly pretty'
			[ɗɛmA2.ɗɛpB2]	
			*[ɗɛpA2.ɗɛpB2]	(tone-coda co-occurrence)
b.	[satB2]	'clean'	[satB2.satB2]	'slightly clean'
			[sanA2.satB2]	
c.	[tɯəB2]	'similar to'	[tɯəB2/A2.tɯəB2]	'slightly similar to'

(6) **C1 words**:

a.	[ɗɔC1]	'red'	[ɗɔC1/A1.ɗɔC1]	'slightly red'
b.	[ɲɔC1]	'small'	[ɲɔC1/A1.ɲɔC1]	'slightly small'
c.	[xwɛC1/A2]	'healthy'	[xwɛC1.xwɛC1]	'slightly healthy'

(7) **C2 words**:

a.	[lwaːŋC2]	'watery'	[lwaːŋC2/A1/A2[1].lwaːŋC2]	'slightly watery'
b.	[ɲaːwC2]	'mushy'	[ɲaːwC2/A1.ɲaːwC2]	'slightly mushy'
c.	[jɛC2]	'easy'	[jɛC2/A1.jɛC2]	'slightly easy'
d.	[kiC2]	'careful'	[kiC2/A1.kiC2]	'slightly careful'

[1]Note that there's an A1 option here as well because there is a preference for A1.C2 combination rather than A2.C2 one, but both are possible regardless, see A.-H. Pham and Pham (2020) for the discussion.

B.2 Pejorative

Monosyllabic base

Noun

(8)	a.	[sɛA1]	'car(s)'	[sɛA1.siəkB1]	'cars (PEJ)'
				*[sɛA1.siəkB2]	
	b.	[ɲaːA2]	'house(s)'	[ɲaːA2.ɲiəkB1/B2]	'houses (PEJ)'
	c.	[ʔaːwB1]	'shirt(s)'	[ʔaːwB1.ʔiəkB1]	'shirts (PEJ)'
				*[ʔaːwB1.ʔiəkB2]	
	d.	[siəkB1]	'circus(es)'	*[siəkB1.siəkB1]	'circusses (PEJ)'
			alternative:	[siəkB1.suːŋ͡mC1]	
	e.	[ɓaːŋB2]	'friend(s)'	[ɓaːŋB2.ɓiəkB1/B2]	'friends (PEJ)'
	f.	[tiəkB2]	'party(s)'	*[tiəkB2.tiəkB1/B2]	'parties (PEJ)'
			alternative:	[tiəkB2.tuːŋ͡mA2]	
	g.	[vʏːC1]	'notebook(s)'	[vʏːC1.viəkB1]	'notebooks (PEJ)'
				*[vʏːC1.viəkB2]	
	h.	[muC2]	'hat(s)'	[muC2.miəkB1/B2]	'hats (PEJ)'

Verb

(9)	a.	[ʔaŋA1]	'to eat'	[ʔaŋA1.ʔiəkB1]	'to eat (PEJ)'
				*[ʔaŋA1.ʔiəkB2]	
	b.	[ɲʏːA2]	'to ask for a favor'	[ɲʏːA2.ɲiəkB1/B2]	'to ask for a favor (PEJ)'
	c.	[ɣɛpB1]	'to couple'	[ɣɛpB1.ɣiəkB1]	'to couple (PEJ)'
				*[ɣɛpB1.ɣiəkB2]	
	d.	[liəkB1]	'to side-eye'	*[liəkB1.liəkB1]	'to side-eye (PEJ)'
				*[liəkB1.liəkB2]	

	e.	[lwiəŋB2]	‘to practice’	[lwiəŋB2.liəkB1/B2]	‘to practice (PEJ)’
	f.	[ɟiəkB2]	‘to destroy’	*[ɟiəkB2.ɟiəkB1]	‘to destroy (PEJ)’
				*[ɟiəkB2.ɟiəkB2]	
	g.	[ŋuC1]	‘to sleep’	[ŋuC1.ŋiəkB1]	‘to sleep (PEJ)’
				*[ŋuC1.ŋiəkB2]	
	h.	[ŋaːC2]	‘to fall’	[ŋaːC2.ŋiəkB1/B2]	‘to fall (PEJ)’

Adjective

(10)	a.	[kaːwA1]	‘tall’	[kaːwA1.kiəkB1]	‘tall (PEJ)’
				*[kaːwA1.kiəkB2]	
	b.	[vaːŋA2]	‘yellow’	[vaːŋA2.viəkB1/B2]	‘yellow (PEJ)’
	c.	[sopB1]	‘porous; fluffy’	[sopB1.siəkB1]	‘fluffy (PEJ)’
				*[sopB1.siəkB2]	
	d.	[satB2]	‘clean’	[satB2.siəkB1/B2]	‘clean (PEJ)’
	e.	[ɲɔC1]	‘small’	[ɲɔC1.ɲiəkB1]	‘small (PEJ)’
				*[ɲɔC1.ɲiəkB2]	
	f.	[lwaːŋC2]	‘watery’	[lwaːŋC2.liəkB1/B2]	‘watery (PEJ)’

Novel words

(11) **Noun**

	a.	[hwaːA1]	‘flower(s)’	[hwaːA1.hwiəkB1/B2]	‘flowers (PEJ)’
	b.	[kwaːA2]	‘present(s)’	[kwaːA2.k(w)iəkB1/B2]	‘presents (PEJ)’
	c.	[ɓenB2]	‘sick’	[ɓenB2.ɓiəkB1/B2]	‘sick (PEJ)’
	d.	[miC2]	‘America’	[miC2.miəkB1/B2]	‘America (PEJ)’

(12) **Verb**

	a.	[mɤːA1]	‘to dream’	[mɤːA1.miəkB1]	‘to dream (PEJ)’
	b.	[sɤːA2]	‘to touch’	[sɤːA2.siəkB1/B2]	‘to touch (PEJ)’

c.	[ɲʏːB1]	‘to miss/remember’	[ɲʏːB1.ɲiəkB1]	‘to miss/remember (PEJ)’
d.	[cajB2]	‘to run’	[cajB2.ciəkB1/B2]	‘to run (PEJ)’
e.	[cʏːC1]	‘to transport’	[cʏːC1.ciəkB1/B2]	‘to transport (PEJ)’
f.	[vɛC2]	‘to draw’	[vɛC2.viəkB1/B2]	‘to draw (PEJ)’

(13) **Adjective**

a.	[sanA1]	‘blue/green’	[sanA1.siəkB1]	‘blue/green (PEJ)’
b.	[hoŋA2]	‘pink’	[hoŋA2.hiəkB1/B2]	‘pink (PEJ)’
c.	[laːwB1]	‘rude’	[laːwB1.liəkB1]	‘rude (PEJ)’
d.	[ɗʏmB2]	‘strong/deep’	[ɗʏmB2.ɗiəkB1/B2]	‘strong/deep (PEJ)’
e.	[hwaːŋC1]	‘in a panic’	[hwaːŋC1.hiəkB1/B2]	‘in a panic (PEJ)’
f.	[ʈeC2]	‘late’	[ʈeC2.ʈiəkB1/B2]	‘late (PEJ)’
g.	[rɔC2]	‘clear’	[rɔC2.riəkB1]	‘clear (PEJ)’

Polysyllabic base

Noun

(14) a. [kaːA2.fɛA1] — [kaːA2.fɛA1 kaːA2.fiəkB1 / *B2]
coffee — coffee $R_{iək}$
‘coffee(s)’ — ‘coffees (PEJ)’

b. * kiəkB2 fɛA1
(number of σ’s in the reduplicative $\neq$ number of σ’s in the base)

c. * kaːA2 fiəkB1 kaːA2 fɛA1
($R_{iək}$ in the second syllable)

d. * kiəkB2 fɛA1 kaːA2 fɛA1
($R_{iək}$ in the first syllable)

e. * kaːA2 fɛA1 kiəkB2 fɛA1
($R_{iək}$ in the third syllable)[2]

f. sɛA1 ɗaːpB2 — sɛA1 ɗaːpB2 sɛA1 ɗiəkB2/B1
vehicle pedal — vehicle pedal vehicle $R_{iək}$
‘bicycle(s)’ — ‘bicycles (PEJ)’

[2]These violations are the same inputs of any lexical categories.

g. [ʔaːwB1 kwɤŋA2]
shirt pant
'clothes'

[ʔaːwB1 kwɤŋA2 ʔaːwB1 kiəkB1/B2]
shirt pant shirt $R_{iək}$
'clothes (PEJ)'

h. cuəŋA2.cuəŋA2
dragonfly
'dragonfly(s)'

cuəŋA2.cuəŋA2 cuəŋA2.ciəkB1/B2
dragonfly $R_{iək}$
'dragonflies (PEJ)'

i. [kɤwA1.lakB2.ɓoB2]
club
'club(s)'

[kɤwA1.lakB2.ɓoB2
club
kɤwA1.lakB2.ɓiəkB2/B1] (Novel data)
$R_{iək}$
'clubs (PEJ)'

Verb

(15) a. laːmA2 ɓaːjA2
do HW
'to do HW'

laːmA2 ɓaːjA2 laːmA2 ɓiəkB2/B1
do HW do $R_{iək}$
'to do HW (PEJ)'

b. * laːmA2 ɓiəkB2 laːmA2 ɓaːjA2
(R in 2nd syllable)

c. * liəkB2 ɓaːjA2 laːmA2 ɓaːjA2
(R in 1st syllable)

d. * laːmA2 ɓaːjA2 liəkB2 ɓaːjA2
(R in 3rd syllable)

e. kɯəjA2 nɔjB1
laugh talk
'to laugh and talk'[3]

kɯəjA2 nɔjB1 kɯəjA2 niəkB1/*B2
laugh talk laugh $R_{iək}$
'to laugh and talk (PEJ)'

f. ŋuC1.ŋeA1
sleep
'to sleep (EMPH)'

ŋuC1.ŋeA1 ŋuC1.ŋiəkB1/*B2
sleep $R_{iək}$
'to sleep (PEJ)'

Adjective

(16) a. vaːŋA2 vaːŋA2
RED yellow
'slightly yellow'

vaːŋA2 vaːŋA2 vaːŋA2.viəkB1/B2
RED yellow $R_{iək}$
'slightly yellow (PEJ)'

[3]in a collective sense

b.	mɤːA1.maːŋA2	mɤːA1.maːŋA2 mɤːA1.miəkB2/B1
	dreamy	dreamy $R_{iək}$
	‘dreamy[4]’	‘dreamy (PEJ)’
c.	sanA1 laːB1	sanA1 laːB1 sanA1.liəkB1
	blue/green leaf	blue/green leaf $R_{iək}$
	‘green’	‘green (PEJ)’
d.	sɨnA1 ɗɛpB2	sɨnA1 ɗɛpB2 sɨnA1.ɗiəkB1
	pretty pretty	pretty pretty $R_{iək}$
	‘pretty’	‘pretty (PEJ)’

Novel words

(17) **Noun**

a.	hɔŋA2 ɗaːB1	hɔŋA2 ɗaːB1 hɔŋA2 ɗiəkB1/B2
	CLF rock	CLF rock CLF $R_{iək}$
	‘rock(s)’	‘rocks (PEJ)’
b.	kɔŋA1 cɔB1	kɔŋA1 kɔB1 cɔŋA1 ciəkB1/B2
	CLF dog	CLF dog CLF $R_{iək}$
	‘dog(s)’	‘dogs (PEJ)’
c.	sɛA1 koB2	sɛA1 koB2 sɛA1 kiəkB2/B1
	vehicle vehicle	vehicles vehicle vehicle $R_{iək}$
	‘vehicles’	‘vehicles (PEJ)’
d.	kajA1 kɔC1	kajA1 kɔC1 kajA1 kiəkB1
	tree grass	tree grass tree $R_{iək}$
	‘vegetation’	‘vegetation (PEJ)’

(18) **Verb**

a.	cajB2 ɲajC1	cajB2 ɲajC1 cajB2 ɲiəkB1/B2
	run jump	run jump run $R_{iək}$
	‘to run and jump’	‘to run and jump (PEJ)’

[4][maːŋA2] doesn’t have meaning in Vietnamese, but it’s an [-aːŋA] reduplicant of an Adj, an unproductive reduplication process

(19) **Adjective**

a. menA1.moŋA1 — menA1.moŋA1 menA1.miəkB1
vast — vast $R_{iək}$
'vast' — 'vast (PEJ)'

b. vojB2.vaːŋA2 — vojB2.vaːŋA2 vojB2.viəkB2/B1
hasty — hasty $R_{iək}$
'hasty' — 'hasty (PEJ)'

c. roŋm͡B2.raːŋA2 — roŋm͡B2.raːŋA2 roŋm͡B2.riəkB2/B1
bustling — bustling $R_{iək}$
'bustling' — 'bustling (PEJ)'

B.3 Exclamation

Noun

(20) a. ʔaːwB1 kwəŋA2 — ʔaːwB1 ʔaːwB1 kwəŋA2 kwəŋA2
shirt pant — shirt shirt pant pant
'clothes' — 'clothes (EMPH)'

b. * ʔaːwB1 kwəŋA2 ʔaːwB1 kwəŋA2 (reduplicative in the form ABAB)

c. kaːA2.fɛA1 — * kaːA2.kaːA2.fɛA1.fɛA1
'coffee' — *Intended*: 'coffee (EMPH)'

d. sɛA1.ɗaːpB2 — * sɛA1.sɛA1.ɗaːpB2.ɗaːpB2
'bicycle' — *Intended*: 'bicycle (EMPH)'

e. cuəŋA2.cuəŋA2 — * cuəŋA2.cuəŋA2.cuəŋA2.cuəŋA2
'dragonfly' — *Intended*: 'dragonfly (EMPH)'

Verb

(21) a. kɯəjA2 nɔjB1 — kɯəjA2 kɯəjA2 nɔjB1 nɔjB1
laugh talk — laugh laugh talk talk
'to laugh and talk' — 'to laugh and talk (EMPH)'

b. * kɯəjA2 nɔjB1 kɯəjA2 nɔjB1 (reduplicative in the form ABAB)

c. ŋuC1.ŋeA1 — * ŋuC1.ŋuC1.ŋeA1.ŋeA1
'sleep' — *Intended*: 'sleep (EMPH)'

d. laːmA2.ɓaːjA2
'to study'

* laːmA2.laːmA2.ɓaːjA2.ɓaːjA2
Intended: 'to study (EMPH)'

Adjective

(22) a. mɤA1.maːŋA2
dreamy
'dreamy'

mɤA1 mɤA1 maːŋA2 maːŋA2
dream dream - -
Intended: 'dreamy (EMPH)'

b. sɨnA1 ɗɛpB2
pretty pretty
'pretty'

sɨnA1 sɨnA1 ɗɛpB2 ɗɛpB2
pretty pretty pretty pretty
Intended: 'pretty (EMPH)'

c. vaːŋA2.vaːŋA2
'yellow (ATT)'

* vaːŋA2.vaːŋA2.vaːŋA2.vaːŋA2
Intended: 'yellow (ATT) (EMPH)'

d. sanA1.laːB1
'green'

* sanA1.sanA1.laːB1.laːB1
Intended: 'green (EMPH)'

(23) a. ɓiəŋB1 ɲaːkB1
lazy lazy
'lazy'

ɓiəŋB1 ɓiəŋB1 ɲaːkB1 ɲaːkB1
lazy lazy lazy lazy
'lazy (EMPH)'

b. hamA1.hɤːC1
'eager'

hamA1.hamA1.hɤːC1.hɤːC1
'eager (EMPH)'

c. ŋɤːA1.ŋaːkB1
'dazed'

ŋɤːA1.ŋɤːA1.ŋaːkB1.ŋaːkB1
'dazed (EMPH)'

d. t^{h}ɤːA1.t^{h}ɤŋC1
'distracted'

t^{h}ɤːA1.t^{h}ɤːA1.t^{h}ɤŋC1.t^{h}ɤŋC1
'distracted (EMPH)'

e. lamA2.liA2
'reticent'

lamA2.lamA2.liA2.liA2
'reticent (EMPH)'

f. ʈuɲ͡mA2.ɗiəpB2
'expansive'

ʈuɲ͡mA2.ʈuɲ͡mA2.ɗiəpB2.ɗiəpB2
'expansive (EMPH)'

g. vojB2.vaːŋA2 — vojB2.vojB2.vaːŋA2.vaːŋA2
‘hasty’ — ‘hasty (EMPH)’

Novel words

(24) **Noun**

a. foB1 fɯəŋA2 — foB1 foB1 fɯəŋA2 fɯəŋA2
street district — street street district district
‘city(s)’ — ‘city(s) (EMPH)’

b. ʔaŋA1 ʔuəŋB1 — ʔaŋA1 ʔaŋA1 ʔuəŋB1 ʔuəŋB1
eat drink — eat eat drink drink
‘eating and drinking’ — ‘eating and drinking (EMPH)’

c. soŋA1 suəjB1 — soŋA1 soŋA1 suəjB1 suəjB1
river stream — river river stream stream
‘rivers and streams’ — ‘rivers and streams (EMPH)’

d. ɗɯəŋA2 saːB1 — ɗɯəŋA2 ɗɯəŋA2 saːB1 saːB1
road road — road road road road
‘roads’ — ‘roads (EMPH)’

e. ɲaːA2 kɯəC1 — ɲaːA2 ɲaːA2 kɯəC1 kɯəC1
house door — house house door door
‘house(s)’ — ‘house(s) (EMPH)’

f. hɔjC1 ɗaːpB1 — hɔjC1 hɔjC1 ɗaːpB1 ɗaːpB1
ask answer — ask ask answer answer
‘Q&A’ — ‘Q&A (EMPH)’

(25) **Verb**

a. ɗɯŋB1 ŋojA2 — ɗɯŋB1 ɗɯŋB1 ŋojA2 ŋojA2
stand sit — stand stand sit sit
‘to be on edge’ — ‘to be on edge (EMPH)’

b. ɣakB2 haːjB1 — ɣakB2 ɣakB2 haːjB1 haːjB1
reap pluck — reap reap pluck pluck
‘to reap; to gain’ — ‘to reap; to gain (EMPH)’

c. ʔaŋA1 ʔiəkB1
eat $R_{iək}$
'eat (PEJ)'

ʔaŋA1 ʔaŋA1 ʔiəkB1 ʔiəkB1
eat eat $R_{iək}$ $R_{iək}$
'eat (PEJ, EMPH)'

(26) **Adjective**

a. maːkB1.mɛC1
cool
'(of weather) cool'[5]

maːkB1 maːkB1 mɛC1 mɛC1
cool cool - -
'(of weather) cool (EMPH)'

b. ɗɛpB2.ɗɛC2
beautiful
'beautiful'[6]

ɗɛpB2 ɗaːA1 ɗɛpB2.ɗɛC2
pretty $R_{aː}$ beautiful
'beautiful (EMPH)'

c. menA1.moŋA1
vast
'vast'

menA1 maːA1 menA1.moŋA1
vast $R_{aː}$ vast

menA1 menA1 moŋA1 moŋA1
vast vast - -
'vast (EMPH)'

d. xwɛC1.xwaŋB1
lively
'lively'[7]

xwɛC1 xaːA1 xwɛC1.xwaŋB1
healthy $R_{aː}$ lively
'lively (EMPH)'

e. cɤmB2.caːpB2
slowly
'slowly'[8]

cɤmB2 caːA2 cɤmB2.caːpB2
slow $R_{aː}$ slowly

cɤmB2 cɤmB2 caːpB2 caːpB2
slow slow - -
'slowly (EMPH)

f. xoA1 xaːŋA1
dry hoarse
'dry/dull'

xoA1 xoA1 xaːŋA1 xaːŋA1
dry dry hoarse hoarse

xoA1 xaːA1 xoA1 xaːŋA1
dry $R_{aː}$ dry hoarse
'dry/dull (EMPH)'

[5][mɛC1] doesn't have meaning in Vietnamese, but it's an [-ɛC] reduplicant of an Adj, an unproductive reduplication process

[6][ɗɛC2] doesn't have meaning in Vietnamese, but it's an [-ɛC] reduplicant of an Adj, an unproductive reduplication process

[7][xwaŋB1] doesn't have meaning in Vietnamese, but it's an [-aŋB] reduplicant of an Adj, an unproductive reduplication process

[8][caːpB2] doesn't have meaning in Vietnamese, but it's an [-aːpB] reduplicant of an Adj, an unproductive reduplication process

g. jeC2.jaːŋA2 — jeC2 jeC2 jaːŋA2 jaːŋA2
easily — easy easy - -
'easily'[9] — jeC2 jaːA2 jeC2.jaːŋA2
easy $R_{aː}$ easily
'easily (EMPH)'

h. sopB1 siəkB1 — sopB1 sopB1 siəkB1 siəkB1
fluffy $R_{iək}$ — fluffy fluffy $R_{iək}$ $R_{iək}$
'fluffy (PEJ)' — 'fluffy (PEJ, EMPH)'

B.4 Intensification (1)

Verb

(27) a. kaŋA2.ɲaŋA2 — kaŋA2.ɲaːA2.kaŋA2.ɲaŋA2
'complain' — 'constantly complain'

b. kamB1.kujB1 — kamB1.kaːA1.kamB1.kujB1
'concentrate' — 'constantly concentrate'

c. lwaːŋB2.cwaːŋB2 — lwaːŋB2.caːA2.lwaːŋB2.cwaːŋB2
'totter' — 'constantly totter'

d. laːmC1.ɲaːmC1 — laːmC1.ɲaːA1.laːmC1.ɲaːmC1
'blabber' — 'constantly blabber'

Adjective

(28) a. mɤːA1.maːŋA2 — mɤːA1 maːA2 mɤːA1.maːŋA2
dreamy — dream $R_{aː}$ dreamy
'dreamy' — 'dreamy (EMPH)'

b. * mɤːA1 maːŋA2 mɤːA1 maːA2
'dreamy (EMPH)'

c. * sɨnA1 ɗaːA2 sɨnA1 ɗɛpB2
'pretty (EMPH)'

d. * vaːŋA2 vaːA2 vaːŋA2 vaːŋA2
'sligthly yellow (EMPH)'

e. * sanA1 laːA1 sanA1 laːB1
'green (EMPH)'

(29) a.

[9] [jaːŋA2] doesn't have meaning in Vietnamese, but it's an [-aːŋA] reduplicant of an Adj, an unproductive reduplication process

lɨnA1.tɨnA1	lɨnA1.taːA1.lɨnA1.tɨnA1
‘nonsensical’	‘nonsensical (EMPH)’

b.

sʏːŋB1.saːkB1	sʏːŋB1.saːA1.sʏːŋB1.saːkB1
‘ill-mannered’	‘ill-mannered (EMPH)’

c.

ɲiB1.ɲanC1	ɲiB1.ɲaA1.ɲiB1.ɲanC1
‘cutesy’	‘cutesy (EMPH)’

d.

ɓiA2.ɓɔmC2	ɓiA2.ɓaːA2.ɓiA2.ɓɔmC2
‘water sound’	‘water sound (EMPH)’

Novel words

(30) **Noun**

a.

ʔaːwB1 ʔiəkB1	ʔaːwB1 ʔaːA1 ʔaːwB1 ʔiəkB1
shirt $R_{iək}$	shirt $R_{aː}$ shirt $R_{iək}$
‘shirts (PEJ)’	‘shirts (PEJ, EMPH)’

b.

ɲaːA2 ɲiəkB1	ɲaːA2 ɲaːA1 ɲaːA2 ɲiəkB1
house $R_{iək}$	house $R_{aː}$ house $R_{iək}$
‘houses (PEJ)’	‘houses (PEJ, EMPH)’

(31) **Verb**

a.

ŋuC1.ŋeA1	ŋuC1 ŋaːA1 ŋuC1.ŋeA1
sleep	sleep $R_{aː}$ sleep
‘to sleep’	‘to sleep (EMPH)’

b.

ɲʏːA2 ɲiəkB1	ɲʏːA2 ɲaːA1 ɲʏːA2 ɲiəkB1
ask.favor $R_{iək}$	ask.favor $R_{aː}$ ask.favor $R_{iək}$
‘ask for a favor (PEJ)’	‘ask for a favor (PEJ, EMPH)’

c.

ɣɛpB1 ɣiəkB1	ɣɛpB1 ɣaːA1 ɣɛpB1 ɣiəkB1
couple $R_{iək}$	couple $R_{aː}$ couple $R_{iək}$
‘to couple (PEJ)’	‘to couple (PEJ, EMPH)’

(32) **Adjective**

a.	maːkB1.mɛC1 cool ‘(of weather) cool’	maːkB1 maːkB1 mɛC1 mɛC1 cool cool - - ‘(of weather) cool (EMPH)’
b.	ɗɛpB2.ɗɛC2 beautiful ‘beautiful’	ɗɛpB2 ɗaːA1 ɗɛpB2.ɗɛC2 pretty $R_{aː}$ beautiful ‘beautiful (EMPH)’
c.	menA1.moŋA1 vast ‘vast’	menA1 maːA1 menA1.moŋA1 vast $R_{aː}$ vast menA1 menA1 moŋA1 moŋA1 vast vast - - ‘vast (EMPH)’
d.	xwɛC1.xwaŋB1 lively ‘lively’	xwɛC1 xaːA1 xwɛC1.xwaŋB1 healthy $R_{aː}$ lively ‘lively (EMPH)’
e.	cɤmB2.caːpB2 slowly ‘slowly’	cɤmB2 caːA2 cɤmB2.caːpB2 slow $R_{aː}$ slowly cɤmB2 cɤmB2 caːpB2 caːpB2 slow slow - - ‘slowly (EMPH)
f.	xoA1 xaːŋA1 dry hoarse ‘dry/dull’	xoA1 xoA1 xaːŋA1 xaːŋA1 dry dry hoarse hoarse xoA1 xaːA1 xoA1 xaːŋA1 dry $R_{aː}$ dry hoarse ‘dry/dull (EMPH)’
g.	jeC2.jaːŋA2 easily ‘easily’[10]	jeC2 jeC2 jaːŋA2 jaːŋA2 easy easy - - jeC2 jaːA2 jeC2.jaːŋA2 easy $R_{aː}$ easily ‘easily (EMPH)’

[10][jaːŋA2] doesn’t have meaning in Vietnamese, but it’s an [-aːŋA] reduplicant of an Adj, an unproductive reduplication process

h.	vaːŋA2 viəkB1	vaːŋA2 vaːA1 vaːŋA2 viəkB1
	yellow $R_{iək}$	yellow $R_{aː}$ yellow $R_{iək}$
	‘yellow (PEJ)’	‘yellow (PEJ, EMPH)’

B.5 Intensification (2) (Triplication)

Monosyllabic base

Adjective

(33)	a.	liA2	liA2.lʏmB2 / *liA2.liA2.liA1/B2
		‘naughty’	‘naughty (EMPH)’
	b.	mɯəkB1	mɯəkB1.mɯəŋA2.mɯəkB2
		‘smooth’	‘smooth (EMPH)’
	c.	saːkB1	saːkB1.saːŋA2.saːkB2
		‘smooth’	‘close to (EMPH)’
	d.	caːkB1	caːkB1.caːŋA2.caːkB2
		‘slapping sound’	‘slapping sound (EMPH)’
	e.	sopB1	sopB1.somA2.sopB2
		‘fluffy’	‘fluffy (EMPH)’
	f.	tuəkB1	tuəkB1.tuəŋA2.tuəkB2
		‘everything’	‘everything & anything’
	g.	tʏkB1	tʏkB1.tʏŋA2.tʏkB2
		‘everything’	‘everything & anything’
	h.	satB2	satB2.sanA2.sanA1
		‘clean’	‘clean (EMPH)’
	i.	tɛwB2	tɛwB2.tɛwA2.tɛwA1
		‘a tiny bit’	‘small (EMPH)’[11]

[11] used in compounds

j. majC1 majC1.majA2.majA1
'a small portion' '(of chance) small (EMPH)'

k. tɯŋC1 tɯŋC1.tɯŋA2.tɯŋA1
'crazy' 'crazy (EMPH)'

Novel words

(34) a. nɤwA1 nɤwC1.nɤwA2.nɤwA1
'brown' *Intended:* 'brown (EMPH)'

b. sanA1 sanC1.sanA2.sanA1
'blue/green' *Intended:* 'blue/green (EMPH)'

c. hɔkB1 hɔkB1.hɔŋA2.hɔkB2
'hot' *Intended:* 'hot (EMPH)'

d. sɤwB1 sɤwB1.sɤwA2.sɤwA1
'ugly' *Intended:* 'ugly (EMPH)'

e. ʔɨtB1 ʔɨtB1.ʔɨnA2.ʔɨtB2
'little' *Intended:* 'little (EMPH)'

f. ɗɛpB2 ɗɛpB2.ɗɛmA2.ɗɛmA1
'pretty' *Intended:* 'pretty (EMPH)'

g. ɓaːkB2 ɓaːkB2.ɓaːŋA2.ɓaːŋA1
'silver' *Intended:* 'silver (EMPH)'

h. ɗɔC1 ɗɔC1.ɗɔA2.ɗɔA1
'red' *Intended:* 'red (EMPH)'

i. ɲɔC1 ɲɔC1.ɲɔA2.ɲɔA1
'small' *Intended:* 'small (EMPH)'

j. ɲaːwC2 ɲaːwC2.ɲaːwA2.ɲaːwA1
'mushy' *Intended:* 'mushy (EMPH)'

k.

lwaːŋC2	lwaːŋC2.lwaːŋA2.lwaːŋA1
‘watery’	*Intended:* ‘watery (EMPH)’

Polysyallbic base

Adjective

(35)

a. xoA1.xoːk͡pB1	xoA1.xoːŋ͡mA1.xoːk͡pB1
‘very dry’	‘very dry (EMPH)’
b. tʰɤːmA1.fɯkB1	tʰɤːmA1.fɯŋA1.fɯkB1
‘very aromatic’	‘very aromatic (EMPH)’
c. satB2.ɓatB1	satB2.ɓanA1.ɓatB1
‘very clean’	‘very clean (EMPH)’
d. tʰɯəA1.tʰɤːkB1	tʰɯəA1.tʰɤːŋA1.tʰɤːkB1
‘very sparse’	‘very sparse (EMPH)’
e. nɔŋ͡mB1.ɓɯŋA2	nɔŋ͡mB1.ɓɯŋA2.ɓɯŋA2
‘flaming hot; feverish’	‘feverish (EMPH)’
f. laŋA1.lok͡pB1	laŋA1.loŋ͡mA1.lok͡pB1
‘rolling around’	‘rolling around and around’
g. ɗɛŋA1.siA2	ɗɛŋA1.siA2.liA2
‘very black’	‘very black (EMPH)’
h. tʰɤmA1.siA2	tʰɤmA1.siA2.liA2
‘very discolored’	‘very discolored (EMPH)’
i. ɗɛŋA1.tʰujA1	ɗɛŋA1.tʰujA2.lujA2
‘burnt black’	‘burnt black (EMPH)’
j. tɨmB1.lɨmB2	tɨmB1.sɨmA2.lɨmB2
‘very purple’	‘very purple (EMPH)’

k.	luA1.ɓuA1	luA1.suA1.ɓuA1
	‘busy’	‘very busy (EMPH)’
l.	lɤːA1.mɤA1	lɤːA1.tɤA1.mɤA1
	‘half-asleep’	‘half-asleep (EMPH)’

B.6 Lexicalized reduplicatives

(36) **(Full) reduplication of nouns (distributive and/or plurality marking**

(Thompson & O’Harrow, 1988; Vu, 2007))

a.	[aːjA1]	‘who’	[aːjA1.aːjA1]	‘everyone’
b.	[ŋɯjA2]	‘person’	[ŋɯjA2.ŋɯjA2]	‘everbody’
c.	[ŋajA2]	‘day’	[ŋajA2.ŋajA2]	‘every day’
d.	[saːŋA1]	‘morning’	[saːŋA1.saːŋA1]	‘every morning’
e.	[ciwA2]	‘afternoon’	[ciwA2.ciwA2]	‘every afternoon’

(37) **(Full) reduplication of adverbs**

a.	[luəŋA1]	‘always; continually’	[luəŋA1.luəŋA1]	‘always; continually’
b.	[hɤːjA1]	‘slightly, somewhat’	[hɤːjA1.hɤːjA1]	‘slightly, somewhat’
c.	[ɗewA2]	‘equal; regular’	[ɗewA2.ɗewA2]	‘equally; regularly’
d.	[tʰɯəŋA2]	‘often’	[tʰɯəŋA2.tʰɯəŋA2]	‘usually’
e.	[maːjC2]	‘for a long time’	[maːjC2.maːjC2]	‘always’

(38) **Onset replacement**

a.	[lanA1]	‘clever’	[lanA1.canA1]	‘hasty’
b.	[ɲaŋA2]	‘eat with front teeth’	[kaŋA2.ɲaŋA2]	‘complain’
c.	[kuəŋB1]	‘panic’	[luəŋB1.kuəŋB1]	‘in a hurry’

(39) **Onset and tone replacement**

a.	[cɤːjA1]	‘play’	[cɤːjA1.ɓɤːjA2]	‘to play around’
b.	[satB2]	‘clean’	[satB2.ɓatB1]	‘completely clean’

(40) **Rhyme replacement**

a.	[kamB1]	'plant'	[kamB1.kujB1]	'concentrate'
b.	[ɓɯkB2]	'upset'	[ɓɯkB2.ɓojB2]	'upset'
c.	[cɤmB2]	'slow'	[cɤmB2.caːpB2]	'slowly'

(41) **Rhyme and tone replacement**

a.	[maːkB1]	'cool'	[maːkB1.mɛC1]	'cool (weather)'
b.	[ɗɛpB2]	'beautiful'	[ɗɛpB2.ɗɛC2]	'beautiful'
c.	[tɨnC1]	'awake'	[tɨnC1.tawB1]	'awake; refreshed'
d.	[jeC2]	'easy'	[jeC2.jaːŋA2]	'easily'

(42) **Cases in which the base and the reduplicant are indistinguishable**

a.	[lɨnA1.tɨnA1]	'nonsensical'
b.	[tʰɨnA1.tʰɨtB2]	'silently'
c.	[ɓiA2.ɓɔmC2]	'water sound'
d.	[sɤːŋB1.saːkB1]	'ill-mannered'
e.	[lɯŋC2.tʰɯŋC2]	'walk slowly and aimlessly'