A UNIFIED ANALYSIS OF SOME VIETNAMESE REDUPLICATION FORMS

Sonny X. Vu
Massachusetts Institute of Technology
<*sonny@sonny.vu*>

1 Introduction[1]

In Vietnamese, one of the main forms of derivational morphology is reduplication. Although reduplicative forms in Vietnamese have been extensively cataloged (Hoàng Văn Hành[2] (1997)) and at least partially described (Emeneau (1951), Thompson (1965), Ngô Thanh Nhàn (1984), Hoàng Văn Hành (1985), Nguyễn Tài Cẩn (1996), Nguyễn Kim Thản (1997), et seq.), they have been subjected to limited analysis (Ngô Thanh Nhàn (1984), Agbayani (1997)) and not examined for general principles. A crucial distinction that often has been overlooked in previous studies is that only a few of the patterns identified as reduplications are productive. In this study, I describe, characterize, and present an OT analysis of four major ***productive*** patterns. I show that rerankings of a core set of well-attested constraints provide a natural account for the data that I present.[3] Most importantly, I show that a number of seemingly disparate properties in these four processes follow from differences in the prosodic structures (Selkirk (1978) et seq.) of the output forms.

The four processes I describe[4] and analyze are: Full Reduplication (FR), reduplication involving Emergence of the Unmarked effects (TETU) (McCarthy and Prince (1994b)), and two processes involving Melodic Overwriting (Alderete, etal. (1997))

---

[1] This paper was supported in part by the Vietnamese Advanced Summer Institute 1997 Program. Special thanks to Nguyễn Thị Minh-Phương for clarifying judgments on all the data in this paper, Cheryl Zoll for guiding this study, and my classmates Calixto Agüero-Bautista and Philippe Schlenker for extensive discussions. Thanks to the participants in the 1997 Workshop in Phonology and Morphology at MIT, the audiences at the 14th Northwest Linguistics Conference and at the 8th Southeast Asian Linguistics Conference, Brian Agbayani, Mark Alves, Cao Xuân H»o, ñinh Lê Thư, Ken Hale, Morris Halle, Michael Kenstowicz, Alec Marantz, Ngô Như Bình, Rob Pensalfini, Phạm Hoa, Shinko Tamura, and Joseph Thai for useful comments and discussions. Usual disclaimers apply.

[2] For the sake of clarity, all Vietnamese names mentioned in this paper will be written in the traditional format and orthography: last, middle, first name. All Vietnamese data will also be presented in the traditional Vietnamese orthography for ease of reference.

[3] Hence, Vietnamese presents a case of conflicting rankings of constraints within a language.

[4] In the data presented, subscripts denote tones (see below for numbering scheme) and superscripts denote stress levels on the following syllable. For example:
^{0}nhà ^{1}toán ^{0}hoc = 'mathematician'
bears the stress pattern 010. The base (BASE) is underlined when it can be determined. Other abbreviations: emph = emphatic; pej = pejorative; att = attenuative; int = intensive; RED = reduplicant. Some of the data in this paper were taken from a number of sources such as Hoàng Văn Hành (1997)'s *Từ Điển Từ Láy* (a dictionary of reduplicative forms produced by the National Linguistics Institute of Vietnam) or from Ngô Thanh Nhàn (1984). All data in this paper are either provided or confirmed by native speaker consultants including, Cao Xuân Hạo, Nguyễn Thị Minh-Phương, Phảm Hoa, and myself.

Mark Alves, Paul Sidwell & David Gil, eds. *SEALSVIII: papers from the 8th meeting of the Southeast Asian Linguistics Society (1998)*. Canberra, Pacific Linguistics, 2007, pp.165-191.

which I will refer to as a-reduplication (a-Red) and iec-reduplication (iec-Red). A summary of some of the differences is provided in Table 1.

FR: sạch → ^{0}sạch ^{1}sạch = 'clean' → 'clean'(+attenuative, informal)[5]
TETU: đẹp → ^{0}đèm ^{1}đẹp = 'pretty' → 'pretty' (+attenuative, formal)
a-Red: ^{1}lò ^{1}mò → ^{0}lò ^{2}mà ^{0}lò ^{1}mò = 'to grope' → 'to grope'(+emphatic)
iec-Red: bán → ^{1}bán ^{1}biệc = 'to sell' → 'to sell' (+pejorative)

Table 1: *Differences between four productive reduplication processes*

Type	Number of syllables in input (BASE)	Stress pattern of output	Syntactic separability of constituents	Avoidance of identical syllables
FR	1	01	no	no
TETU	1	01	no	yes
a-Red	2	0201	no	yes
iec-Red	1 or 2	11, 0101	yes, yes	yes

In Table 1, 'syntactic separability' refers to the observation that some of the reduplicative forms in Vietnamese can have syntactic material inserted between BASE and RED if they can be determined. The last column in Table 1 refers to the fact that some processes do not allow the copying of identical material while others do. A shared feature of the four processes is the fact that the number of syllables in the base is equal to the number of syllables in the RED. Thus, for example, the output of a-Red always has four syllables whereas the output for FR is always disyllabic since a-Red always takes a disyllabic base and FR a monosyllabic base. Another shared feature is that all four processes preserve the grammatical category of BASE.

In addition to exploring the differences between these four processes, I also examine the common features shared among them. These include: complete preservation of the base in the input, lack of double reduplication (the application of a reduplicative process on a reduplicated form), and respect of regular tone-coda co-occurrence restrictions that appear elsewhere in the language.

One of the main empirical observations made in this study is the distinction between productive and non-productive reduplication patterns. Although both types obey some of the constraints I posit, only the productive patterns are assumed to be produced by the grammar, whereas the non-productive cases (the majority of reduplicative forms in Vietnamese) are not. Although this distinction has been alluded to in the Vietnamese linguistics literature (e.g., Ngô Thanh Nhàn (1984)), many of the descriptions of Vietnamese reduplication (e.g., Nguyễn Tài Cẩn (1996), Nguyễn Kim Thản (1997)) assume that all reduplicative forms are formed in the grammar by various rules; however, a cursory review of the data shows that these rules over-generate the possible forms in the language.

[5] All final obstruents in Vietnamese are unreleased. I do not indicate this in the data for typographical simplicity.

2 Vietnamese Phonology and Theoretical Assumptions

2.1 *Relevant Vietnamese Phonology*

Tones: In most Northern dialects, the dialects of focus in this study, Vietnamese has six phonologically contrastive tones. Every syllable is associated with one of these tones. Although I will use the same notation as Burton (1992)'s, I depart from his categorization by viewing the tone pairs {1, 4}, {2, 5}, and {3, 6} to be related not by contour shape (since they are clearly not similar in shape, e.g., tone 2 is rising whereas tone 5 is dipping) but rather, in terms of their 'bundles' of phonetic features.[6] See Wannamacher (1997) for a related discussion on voice quality as a potential tonal feature. Tones 1 and 4 are both lax in terms of glottal stricture and medium in terms of length (for more discussion of length as a tonal parameter, see Alves (1997b)). Tones 2 and 5 have tense glottal stricture, are glottalized at the end, and are short. Tones 3 and 6 are also both tense but are long, in fact, longer than all the other tones. Perhaps one way to characterize tonal markedness in Vietnamese is to say that tones that are lax are unmarked and the ones that are not lax are more marked. For a detailed phonetic description of Vietnamese tones, see Han (1969). See Edmondson and Lợi (1997) for a detailed phonetic study of the tones in Northern Vietnamese. The main motivations for the coupling of tones as shown in Table 2 are: their synchronic phonological correlations (e.g., their related distributions in reduplication as described in this paper, see Võ Xuân Hạo (1997) for a detailed discussion) and their tonogenetic history.[7]

Table 2: *Vietnamese tone registers and contours*

		Contour[8]		
		l	**h**	**lh**
Register	+High	1	2	3
	-High	4	5	6

There is no phonological tone sandhi except in a few domains of the grammar, e.g., reduplication. This characterization is motivated not only in terms of their phonological 'well-behavedness' but also by a number of tonogenetic accounts such as Haudricourt's hypothesis (Haudricourt 1954) of tonal splitting due to the presence of voiceless obstruent codas lowering the pitch of tones in early Vietnamese.

---

[6] Rolf Noyer has mentioned in personal communication that perhaps the characterization based on the contour and register parameters can be maintained. The actual contour shape of tone 5 might then just be a phonetic reflex of the tone being in the lower register and having the glottalization at the end.

[7] Historically, Vietnamese developed phonetically conditioned allophonic features in different syllable types, A (open syllables), B (syllables with final fricatives), and C (syllables with final stops p/t/c/k/?). As the coda consonants in Vietnamese changed (fricatives and glottal stops, which are well attested in other Mon-Khmer languages, were lost in Vietnamese), those phonetic categories became phonemic (i.e., unpredictable based on environment). The next step was the conditioning of height based on the voicing of the initial. Tonal categories A, B, and C each developed two allophones, with voiceless onsets resulting in higher tones and voiced, lower tones. Eventually, Vietnamese initials changed, thereby masking the original voicing, but leaving the tonal height, which became phonologically contrastive. Thus, 3 tonal phonemes became 6. (Haudricourt (1954), Alves (1997b))

[8] Or better put, grouping with respect to voice quality.

In terms of tonal features, I adopt Agbayani's correlation of [α vocal fold] features with tone contours on vowels and voice distinctions on consonants:

[stiff] ⇒ C[-voice] or V[**h**] and [slack] ⇒ C[+voice] or V[**l**]

Vowels: Vietnamese has nine vowels. Although vowels obligatorily lengthen in open syllables, two of the vowels /ɤ/ and /ɑ/ are contrastive in terms of length in other environments as well. The place features of the vowels are the most relevant to the present discussion. For a detailed phonetic study of Vietnamese vowels, see Trần (1967) and Han (1966).

Vowel:	[i]	[e]	[ɛ]	[u]	[o]	[ɔ]	[ɯ]	[ɤ]	[ɑ]
Place:	+cor	+cor	+cor	+lab +dor	+lab +dor	+lab +dor	+dor	+dor	+rad

Tone-coda co-occurrence restriction: Syllables with final obstruents, i.e., [p, t, k], can only be assigned tones with contour value **h**, i.e., tones 2 and 5. Otherwise, any of the six tones can be assigned to any other syllable type. For typographic reasons, diphthongs and triphthongs are represented here the two and three vowels, respectively.

Other: Vietnamese lacks consonant clusters and has only a limited number of possible codas. The possible syllable types are: C(G)V(ː)C and C(G)Vː where G is a glide. Note that C(G)V is not allowed, i.e., syllables with only an onset (without or without a glide) and a short vowel. For typographic reasons, diphthongs and triphthongs are represented here the two and three vowels, respectively.

Although any consonant from the phonetic stock can be an onset, the only possible codas are the voiceless unreleased obstruents [p, t, k], nasals [n, m, ŋ] and glides [w, j]. There are also no co-occurrence restrictions on vowels (or vowel clusters) and onsets. The most noticeable phonotactic effects are the co-occurrence restrictions of certain vowels and coda consonants and the co-occurrence restriction on certain tones and codas. The matching of alveolar and palatal place features between main vowels and coda consonants is the most salient pattern of the former: the vowels [i.e., ɛ] cannot occur with the velar codas [k, ŋ].[9] For a detailed discussion of the general nature of the Vietnamese phonological system, see Thompson (1965) and Đoàn Thiện Thuật (1977).

2.2 *Theoretical Assumptions*

Optimality Theory: My analysis is presented in the Optimality Theory (OT) framework in which universal but violable constraints are posited. Within this theory, languages differ according to the way these constraints are ranked with respect to one another. The main mechanisms in OT are:

> *Gen:* function which generates all the possible candidates for a given input
> *Eval:* function which evaluates the optimality of each of the candidates with respect to a language-specific constraint ranking for the given input

[9] Pham (1997) gives an account of the distribution of final consonants in 4 dialects (Hanoi, Saigon, Hue, Quangnam) and for the co-occurrence restriction of velars after front vowels.

Con: set of universal constraints

The optimal candidate selected by Eval is the one which violates the highest ranking constraints (taken from Con) the least number of times, i.e., the one which satisfies the most number of constraints starting from the most highly ranked one. In the case where candidates tie, in the sense that the highest ranking constraint they violate is the same one, the candidate which violates this constraint the least number of times wins; if they tie on this measure, then the one which satisfies the next highest constraint the least number of times wins, and so on.

One notion in OT which is crucial to my analysis is Correspondence Theory (CT) (McCarthy and Prince (1994a), McCarthy and Prince (1995)), an approach which provides a mechanism for the general purpose of string comparison. A correspondence relationship is said to measure the *faithfulness* of the output to the input, i.e., how much deviation exists between the output and the original input form. In the theory of reduplication, this notion of Correspondence is also useful in the comparison of how different the reduplicant (RED) is from the base (BASE) as well as how faithful the output is to the input. The correspondence relation between two strings is defined in 0.

Correspondence: Given two strings S_1 and S_2, correspondence is a relation $\Re$ from the elements of S_1 to those of S_2. Segments $\alpha \in S_1$ and $\beta \in S_2$ are referred to as correspondents of one another when $\alpha \Re \beta$. (McCarthy and Prince (1995))

Prosodic structure: In this study, I rely heavily on making reference to prosodic structure. I posit these structures as a part of my analysis because a number of generalizations, such as stress patterns and syntactic separability, of these processes can be accounted for purely in prosodic structural terms. In following this approach, I have assumed the Prosodic Hierarchy Hypothesis (Selkirk (1978), Nespor and Vogel (1982), Hayes (1989), et seq). According to this theory, prosodic domains are divided into a number of salient levels on which various phonological and morphological processes can take place. The hierarchy is given in 0:

Prosodic Hierarchy	Domain
ϕ	Phonological Phrase
|	
ω	Phonological Word (or Prosodic Word)
|	
Ft	Foot
|	
σ	Syllable
|	
μ	Mora

The levels that are particularly relevant to this discussion are the foot and the phonological word. In Vietnamese, as in many languages, each foot is assigned one stress. It is assumed that in Vietnamese, this level ordering is strict and exhaustive in the sense that there are no other levels intervening between any two given levels and that for a given utterance all of the levels exist in the prosodic structural representation. Therefore, for example, given a phonological word, one cannot insert another phonological word inside

of it. However, within a phonological phrase, a phonological word may be inserted in between two other words. This last fact will be the basis of one of my arguments for different processes having different prosodic structures.

Reduplication: In this paper, I assume ideas from Generalized Template Theory, an approach couched in the OT framework that views the shape of reduplicants as the result of morphological specifications. This approach views reduplication as a process of concatenating a RED affix that is phonetically empty in the underlying representation which is later filled in by various phonological and/or morphological processes. (McCarthy and Prince (1994a), Urbanczyk (1997)) Proceeding along these line, Alderete et. al. (1996) makes a distinction between two different kinds of processes which provide the reduplicant with phonetic content. Both processes involve the insertion of fixed material; they are characterized as: TETU processes and Melodic Overwriting (henceforth, MO) (See Yip (1992)). TETU arises from phonological constraints on markedness (segmental and tonal) whereas MO derives from morphological sources. There are examples of both kinds of processes in Vietnamese.

3 Data: Descriptions and Generalizations

3.1 *Productive processes*

In this section, I describe in detail the four types of reduplication processes that I will be investigating in this paper. The reason for selecting these processes is that they are the most productive, if not the only ones, in the language. As shown below, many of the forms which do not fall in any of these four processes should be treated as lexicalized forms. Independent evidence for this proposal includes the fact that they have irregular changes in meaning as well as non-productive, non-predicable phonological and morphological form, and that they can themselves undergo reduplication.

Full reduplication: One of the most common types of reduplication in Vietnamese is full or total reduplication. In these cases, RED is an exact copy of BASE both segmentally and tonally. The semantic change in these cases is usually attenuative. There are a number of cases given in Thompson (1965) which report that plural and distributive meanings are possible when nouns undergo FR as seen in examples 0; however, these cases only exist in certain dialects and are very unnatural in many dialects, including most of the Northern dialects. In general, FR can only operate on stems that are adjectives (a-d, i, j), adverbs (g, h), or verbs (e, f). The stress patterns that obtain in these forms are regularly 01 (personal judgment, Cao Xuân Hạo (1998a)). As shown below, in terms of meaning changes and stress assignment, these forms most closely resemble the TETU forms mentioned above.

Full reduplication examples

a.	nhanh → ^{0}nhanh ^{1}nhanh	(+H,**l**)(+H,**l**)	fast (+att)
b.	đẹp → ^{0} đẹp ^{1} đẹp	(-H,**h**)(-H,**h**)	pretty (+att)
c.	cười → ^{0} cười ^{1} cười	(-H,**l**)(-H,**l**)	smile (+att)
d.	nhỏ → ^{0} nhỏ ^{1} nhỏ	(+H,**lh**)(+H,**lh**)	small (+att)
e.	sách → ^{0}sách ^{1}sách	(+H,**h**)(+H,**h**)	book (+dist)

These forms do not have the separability effects. Also, they cannot form the base for any reduplication process (or any other derivational morphological process in general, e.g., nominalization) and only monosyllabic words can serve as the base for FR. (Exceptions are usually more literary forms.)

Reduplication via replacement with unmarked material: As mentioned above, a significant number of reduplication cases in Vietnamese can be accounted for by the claim that RED is filled with material that is unmarked where possible and is minimally different from the base. In this way, these cases are productively formed by the grammar. The examples in 0 - 0 (data taken and readjusted[10] from Agbayani (1997)) illustrate the proposal that [+stiff vocal cord] as well as [+labial] and [+dorsal] vowel place features (as opposed to [+coronal] and [+radical]) are marked structures in Vietnamese. The consequence of these marked structures is detailed in and 0 and 0 and illustrated in examples 0 - 0. Independent evidence for the marked status of these forms include the relative sparseness of these forms indicated in various statistical studies of text corpora (Ngô Thanh Nhàn (1984), Võ Xuân Hạo (1997)) and the forms' relative non-occurrence in loan phonology as shown in studies such as Nguyễn Đinh-Hòa (1980).

Consequences of tonal markedness
The marked tonal feature [+stiff vocal cord] is associated with tones 2, 3, 5, and 6, thus making these tones more marked than the ones that do not have this feature, i.e., tones 1 and 4.

Consequences of vowel place markedness
The markedness of vowel place features [+labial] and [+dorsal] results in the vowels [u, o, ɔ] being the most marked for having both of these, followed by [i, e, ɛ] for having only one of these features, and finally by the least unmarked vowels [ɯ, ɤ, ɑ] for having neither of these features. This is deduced from the vocalic stock whose features are summarized in section 2.

To illustrate how unmarked material emerges in reduplication, I present a set of data exemplifying the regular tonal changes that occur in TETU reduplication. The examples below differ minimally on their basic syllabic structure: syllables with no coda in the base 0, nasal codas 0, and voiceless obstruent codas 0. See Agbayani (1997) for a more detailed discussion of marked (and unmarked) structures in Vietnamese.

In examples 0 and 0, the register of the tone associated with the base is preserved in RED and the contour of the tone in RED changes to l. Facts concerning tone harmony were also observed in Vũ Thế Thạch (1994). Note that in the case where BASE already has an unmarked tone, the tone does not change in RED.

Syllable type: no coda in BASE
(-H, c) in BASE → (-H, l) in RED; no segmental changes

i.	nhiều	(-H**,l**)	be a large amount
ii.	^{0}nhiều ^{l}nhiêu	(-H**,l**)(-H**,l**)	be a rather large amount

[10] Glosses readjusted based on native speaker judgments; stress patterns added.

iii.	nhẹ	(-H,**h**)	be light (in weight)
iv.	^{0}nhe ^{1}nhẹ	(-H,**l**)(-H,**h**)	be rather light

(+H, c) in BASE → (+H,**l**) in RED; no segmental changes

i.	hiu	(+H,**l**)	(breeze) is gentle
ii.	^{0}hiu ^{1}hiu	(+H,**l**)(+H,**l**)	blow very lightly
iii.	khá	(+H,**h**)	rather good
iv.	^{0}kha ^{1}khá	(+H,**l**)(+H,**h**)	rather mediocre
v.	nhỏ	(+H,**lh**)	small
vi.	^{0}nho ^{1}nhỏ	(+H,**l**)(+H,**lh**)	rather small

(8) Syllable type: sonorant (nasal) coda in BASE

(-H,c) in BASE → (-H,**l**) in RED; no segmental changes

i.	buồn	(-H,**l**)	be sad
ii.	^{0}buồn ^{1}buồn	(-H,**l**)(-H,**l**)	be a little sad
iii.	chậm	(-H,**h**)	be slow
iv.	^{0}châm ^{1}chậm	(-H,**l**)(-H,**h**)	be somewhat slow
v.	loãng	(-H,**lh**)	diluted
vi.	^{0}loãng ^{1}loãng	(-H,**l**)(-H,**lh**)	rather diluted

(+H,c) in base → (+H,**l**) in reduplicant; no segmental changes

i.	xanh	(+H,**l**)	be blue, green
ii.	^{0}xanh ^{1}xanh	(+H,**l**)(+H,**l**)	be bluish, greenish
iii.	trắng	(+H,**h**)	be white
iv.	^{0}trăng ^{1}trắng	(+H,**l**)(+H,**h**)	be whitish
v.	ẩm	(+H,**lh**)	be humid
vi.	^{0}âm ^{1}ẩm	(+H,**l**)(+H,**lh**)	be slightly humid

In 0, note that the same tonal change occurs as seen in the above examples. However, in order to obey the tone-coda co-occurrence restriction, the change also drives a segmental change when the base has an obstruent in the coda: the coda in the reduplicant is the homorganic nasal corresponding to the obstruent in the coda of the base.

Syllable type: obstruent coda in base

(α**H**,**h**) in BASE → (α**H**,**l**) in RED; obs]$_{\sigma}$ in base → nasal]$_{\sigma}$ in RED

i.	đẹp	(-H,**h**)	be beautiful
ii.	^{0}đèm ^{1}đẹp	(-H,**l**)(-H,**h**)	be rather pretty
iii.	sạch	(-H,**h**)	be clean
iv.	^{0}sành ^{1}sạch	(-H,**l**)(-H,**h**)	be rather clean
v.	tốt	(+H,**h**)	be good, fine
vi.	^{0}tôn ^{1}tốt	(+H,**l**)(+H,**h**)	be rather good
vii.	chắc	(+H,**h**)	certain
viii.	^{0}chăng ^{1}chắc	(+H,**l**)(+H,**h**)	more or less certain

Note that there are cases where a base can be reduplicated via FR as well as TETU. The difference in these cases would be the reduplicative forms' formal register. For example, although both đèm đẹp and đẹp đẹp are possible reduplications of đẹp, the former is more formal than the latter. This difference can also be understood in terms of a colloquial versus literary distinction where the FR form is the more colloquial of the two.

Reduplication via Melodic Overwriting I: iec-reduplication: One of two productive forms of reduplication in which fixed segments appear in the reduplicant is iec-Red (henceforth,

RED_{iec}). This suffixing process adds a pejorative meaning to base. 0 and 0 provides a summary of the generalizations on the form and distribution of this process with illustrative examples in 0.

(10) Examples of iec-Red forms

a.	[1]bạn [1]biệc	(-H,**h**)(-H,**h**)	friends (+pej)
b.	[1]ngủ [1]ngiếc	(+H,**lh**)(+H,**h**)	to sleep (+pej)
c.	[0]uống [1]thuốc [0]uống [1]thiếc	(+H,**h**)(+H,**h**)(+H,**h**)(+H,**h**)	to take medicine (+pej)
d.	*[1]liếc [1]liếc	(+H,**h**)(+H,**h**)	to scam (+pej)[11]
e.	*[1]tiếc [1]tiếc	(+H,**h**)(+H,**h**)	to regret (+pej)[12]

(11) Generalizations on the form of iec-Red

a. RED_{iec} can only be attached at the right edge of BASE.

b. RED_{iec} is identical to BASE except the last rhyme of RED_{iec} must be iec with tone (+H,**h**). In a number of Southern and Central dialects, the register of the tone that is associated with the rhyme replacement matches the register of the corresponding rhyme in BASE.

c. The stress pattern for disyllabics formed by this process is 11; the pattern for quadrasyllabics is 0101.

d. This process cannot operate on BASE with rhyme = iec with tone (αH,**h**). In these cases, the form simply does not exist, nor is there any regular alternative form. The meaning would be conveyed either paraphrastically or by the use of a replacement lexical form (see the starred examples in 0). This is in contrast to analogous cases in Abkhaz (m-reduplication; see Bruening (1996)) and Telegu (p-reduplication; see Alderete et. al. (1997)) where repetition is avoided by the use of a replacement allomorphic form. For example, in Abkhaz, for bases which have *m-* in the onset, an m-reduplication process is replaced by a p-reduplication process to avoid identical copying.

(12) Generalizations on the distribution of iec-Red

a. The process can occur on a base of any lexical grammatical category (noun, verb, adjective, etc.) of one or two syllables. In the case of disyllabic verbs, it can only apply to non-simple verbs (verb-verb or verb-object compounds) or borrowed disyllabic verbs.

b. There may be additional material intervening between BASE and RED of a reduplicative form that was produced by iec-Red. A discussion of the constraints on what kinds of syntactic structures can be inserted is beyond the scope of this paper; however, a number of illustrative examples are provided in 0.

c. In general, iec-Red can only apply to bases with a maximum of two syllables.

In general, in a significant number of reduplicative forms in Vietnamese BASE and RED may be separated from each other and syntactic material can be inserted. The only process out of the four considered which creates forms that have this effect is iec-Red.

---

[11] There is no replacement form for this; the meaning would have to be expressed paraphrastically.

[12] [1]tiêng [1]tiec would be a replacement form.

Other lexicalized reduplicative forms have this property as well.[13] Below, I account for the fact that iec-Red forms exhibit these effects while a-Red ones do not.

(13) Examples of separability of iec-Red forms

a. Con chẳng chịu học hành gì cả cứ ban với biếc suốt ngày.
you NEG try study what all just friends and RED_{iec} all day
"You never study or do anything, you're just with your friends all day."

b. Tại sao con cứ khóc hoài vậy không ngủ không ngiếc bao giờ.
why you just cry all the time Q NEG sleep NEG RED_{iec} ever
"Why are you crying all the time? You're never sleeping or anything!"

c. Anh ấy chẳng may áo mà cũng chẳng may iếc gì cả.
he that NEG sew shirt and also NEG RED_{iec} what all.
"He doesn't sew shirts or anything at all."

Reduplication via Melodic Overwriting II: a-reduplication: Another regular process of Melodic Overwriting in Vietnamese is the replacement of a coda rhyme by *-a:* in RED in the reduplication of some disyllabic verbs. I refer to the reduplicant in this process as RED_a. 0 presents a summary of the generalizations about the forms created by this process. Examples in 0 are representative of forms produced by this process.

(14) Example of a-Red forms

a. rộn ràng → ^{0}rộn ^{2}rà ^{0}rộn ^{1}ràng = to be tumultuous (+emph)
(-H,**h**)(-H,**l**)(-H,**h**)(-H,**l**)

b. may mặc → *^{0}may ^{2}mà ^{0}may ^{1}mặc = to sew and to wear
(+H,**l**)(-H,**l**)(+H,**l**)(-H,**h**)

c. ăn cơm → * ăn cơm ca ăn cơm = to have a meal[14]
(+H,**l**)(+H,**l**)(+H,**l**)(+H,**l**)

(15) Generalizations on the form of a-Red

a. RED_a can only attach to the left of BASE.

b. All of BASE is copied exactly (segments as well as tones), except the rhyme of the second syllable in RED has the following form:
   i. vowel: always *a*[15] and no coda
   ii. tone: register is the same as the register of the corresponding tone in the base; the contour of the tone must be **l**.

c. The stress pattern in the quadrasyllabic output is: 0201

The process can only occur on bases that are simple verbs (i.e., verbs that are not a compound of two verbs or an incorporation of an object and a verb) that are of two syllables. Thus, a-Red targets much more specific bases than iec-Red. Examples 0 show

---

[13] Compounds are also analogous to these i‰c-Red and lexicalized forms in that, in addition to the stress properties mentioned above, they also have this separability effect.

[14] Literally, 'eat' + 'rice'. Can also have the meaning 'to eat rice.'

[15] Here, we see an Emergence of the Unmarked effect within a Melodic Overwriting scheme. Also, there are a number of cases of ɤ: (back, middle, unrounded) being an acceptable rhyme, but there only a very limited number of these cases (~3-4). In any case, ɤ: is fairly unmarked so an extension on the Emergence of the Unmarked argument to vowels could account for this.

that a-Red is not possible on serial or compound verbs and verb-object incorporations. The process cannot be applied to forms in which the second syllable has rhyme *-a*. As in the similar exceptional case for iec-Red, the meaning of the avoided form is expressed either paraphrastically or is replaced by another lexical form with the same meaning. This is seen in examples 0. Forms produced by a-Red do not exhibit any separability effects as seen in compounds, iec-Red forms, and other polysyllabic forms.

(16) Examples of a-Red in which certain outputs are prohibited
a. *lê la lê la (lê la = to wander drunkenly)
b. lê lê la la = to wander drunkenly (+emph, +crit)

3.2 *Non-productive (Lexicalized) Forms*
One of the main empirical observations that I make in this paper is that many of the forms that traditionally have been thought to be the result of reduplication processes are actually lexicalized forms. By "lexicalized," I am referring to the fact that a form has an arbitrary form, an arbitrary meaning, or both, and that it is not actively produced by grammatical processes (phonological, morphological, etc.). Thus, the motivation to refer to these cases as patterns may have been because many of the forms do indeed pattern together in that there are distinct sets of cases where only the onset is reduplicated, only the onset and the tone is reduplicated, etc. However, the strongest evidence for the idea that these forms are lexicalized is that they are all non-productive in their phonological form and/or their meaning. The choice of whether the onset, coda, tone, or some combination thereof is preserved is not predictable from the form of (what would be) the 'base.'

For cases in which a form exhibits reduplicative phonology in some way, i.e., some salient subpart of the constituent morphemes resemble each other, it may simply be that the form was produced by some sort of reduplicative process in the past but has changed over time both in form and in meaning, and now behaves just like any other polysyllabic word. In cases where neither of the morphemes in the form are free morphemes, it seems even clearer that the form was not generated by the grammar but rather stored in memory as a single lexical entry.

These lexicalized forms exhibit many of the same properties, listed in 0, as compounds do, and hence are similar to iec-Red forms.

(17) Shared properties of lexicalized reduplicatives and compounds
a. Stress pattern is 11.
b. Exhibit separability effects.

Note that the examples listed below exhibit nearly every possible type of partial reduplication given the assumption that the syllable is the minimal unit of analysis. (Ngô Thanh Nhàn (1984))

Examples of different types of lexicalized reduplicative forms

(18) Onset replacement:
a. rối (to entangle) → bối rối (to embarrass)
b. lanh (clever, quick) → lanh chanh (hasty, hurried)

(19) Rhyme and tone replacement:
a. rộn (busy) → rộn ràng (happily busy)
b. mò (to grope) → lò mò (to grope for a long time)

(20) Onset and tone replacement:
a. chơi (to play) → chơi bời (to play mischievously)
b. tấp (to rub up against) → tấp nập (to be crowded)

(21) Rhyme replacement:
a. cặm (to plant into) → cặm cụ (to concentrate)
b. tỉnh (refreshed) → tỉnh táo (awake, not sleepy)

(22) Nuclear vowel replacement:
a. mập (fat) → mập mạp (fat (only concerning children))
b. rao (to advertise) → rêu rao (to be obnoxiously loud)

(23) Coda replacement:
a. xanh (blue) → xanh xao (pale)
b. chắc (solid) → chắc chắn (sure, certain)

(24) Cases where BASE and RED are not discernible:
a. lững thững = to walk slowly and aimlessly
b. lay láy = to sparkle

3.3 *Summary*
The similarities between the four processes are list below. Differences between the individual processes are summarized in Table 1.

- BASE is always preserved (segmentally and tonally) in the output.
- The number of syllables in RED is equal to that in BASE.
- The forms all respect the tone-coda co-occurrence restriction.
- There can be no double reduplications in a productive sense.

4 Analysis

I first show that a number of highly-ranked constraints are active in all four processes (including TETU as partially analyzed in Agbayani (1997)). This accounts for the similarities. Next, I show that two constraints are active in all four processes, but are ranked differently depending on the process. For each process, I present arguments for the relative ranking of these constraints. I do this by first positing a prosodic structure for the outputted reduplicative form, and then illustrating that the properties mentioned above follow from these structures by computing a number of example tables based on the constraints proposed.

I am assuming that RED is phonetically empty in the input and is filled in via correspondence relations which attempt to map as many similar segments as possible from

the base without the insertion of additional material. I posit prosodic structure for candidates provided by GEN but not for the underlying representations. EVAL determines the optimal candidate based on segmental, tonal and prosodic content with respect to the constraints from CON.

The analysis naturally follows from the observation that the surface properties of the different processes vary according to their prosodic structures. For FR and TETU processes, since RED is an affix and is only one syllable long, under the assumption that the Prosodic Hierarchy is strictly layered, the prosodic structures of the output forms follow from alignment and foot-binarity constraints. In these two processes, the structure of the output is a prosodic word with one foot consisting of two syllables as shown in 0. Similarly, for a-Red, RED is an affix with two syllables, and so the output structure, a prosodic word with two disyllabic feet, is predicted from the same alignment and foot-binarity constraints. In these three processes, properties such as differential stress patterns (i.e., 01 in TFR and TETU, 0201 in a-Red) within a single word and the non-separability of the constituent morphemes follow from the proposed prosodic structure.

In the case of iec-Red, non-differential stress patterns (i.e., 11 and 0101 for mono- and disyllabic bases, respectively) and the syntactic separability of the base and the reduplicant support the hypothesis that RED is itself a prosodic word (structure given in 0).

I use an Optimality Theoretic approach via Correspondence Theory because it can provide a natural account of the properties of all four processes, thereby giving a unified analysis of a large portion of the (productive) Vietnamese reduplication system.

4.1 *Highly-ranked (Undominated) Constraints*

I posit the constraints listed below to be active and highly-ranked in all four processes. In fact, these constraints are undominated except for the fact that there is one process-specific constraint (alignment of RED_{iec}) which is ranked above these constraints in iec-Red. These constraints account for the prosodic structures in the output forms, whereas the relatively ranked constraints posited in the following sections account for the segmental and tonal content of the forms.

(25) Alignment constraints

ALIGN: Abbreviation for the following three alignment constraints which determine the position of RED with respect to BASE and which level RED combines with BASE: (McCarthy and Prince (1993))

a. **ALIGN-L (RED, ω)**: For every RED affix, there is a ω such that the left edge of RED coincides with the left edge of ω.

b. **ALIGN-R (ω, Stem)**: For every ω there is a stem (root + affix) (McCarthy and Prince 1993, section 4) such that the right edge of ω coincides with the right edge of stem. This forces the constituents in FR, TETU, and a-Red to be within the same stem and therefore not separable by anything that is larger than a root, such as a stem or a morphological word, in the Morphological Hierarchy.[16]

[16] Here, I assume the Morphological Hierarchy to be the traditional one as assumed in (McCarthy and Prince (1994a)):
Morphological Word → Stem; Stem → Stem Affix; Stem → Root

c. **ALIGN-L (ω, H(Ft))**: For every ω there is a head foot H(Ft) such that the left edge of ω coincides with the left edge of H(Ft).

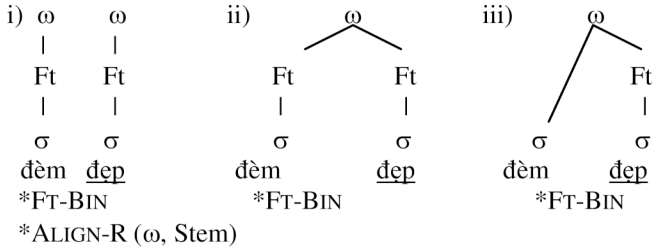
ALIGN-L (RED, ω) prevents cases like *<u>*lếch kếch*</u> *lếch ka* where RED appears on the wrong side of the ω. ALIGN-L (ω, H(Ft)) determines the placement of the head foot which accounts for the fact that primary stress falls on the first foot and not the second one in a-Red output forms.

(26) Foot Binarity
FT-BIN (σ): All feet are binary under syllabic analysis.

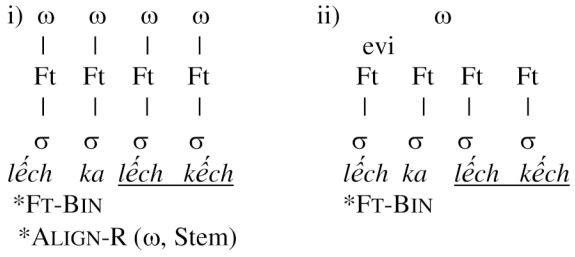
ALIGN-R (ω, Stem) and FT-BIN prevents cases like 0 where the output has a different prosodic structure, which would result in inaccurate predictions for the various properties of the output. Ft-Bin rules out all of the structures in 0 except b),iv) which is ruled out by ALIGN-R (ω, Stem). The descriptive facts about differences in the prosodic structure of the different forms are accounted for by alignment and binarity constraints.

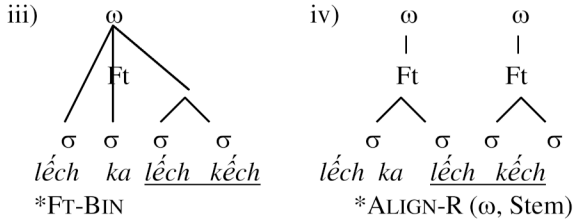
(27) Ill-formed prosodic structures; constraint(s) violated

a. đèm <u>đẹp</u> = pretty (+att)



b. *lếch ka* <u>*lếch kếch*</u> (+H,**h**)(+H,**l**)(+H,**h**)(+H,**h**) to clatter (+emph)





(28) Input-Output Faithfulness

IO-FAITH: The base in the input must not be altered in the output. Technically, this is actually three constraints in one: IO-DEP, IO-MAX, and IO-IDENT, which respectively prohibit insertion, deletion, and non-identity of corresponding segments in the output. (McCarthy and Prince (1993))

This constraint prevents BASE from changing. Therefore, if there is ever a mismatch between BASE and RED, it is not because BASE changed, but rather it is because of partial copying or the insertion of fixed and/or unmarked material. This is why examples such as b) below are good rather than c), which would otherwise have fewer marked features overall.

(29) Example: BASE is always preserved

a.	đẹp	(-H,**h**)	pretty
b.	đèm đẹp	(-H,**l**)(-H,**h**)	pretty (+att)
c.	*đèm đèm	(-H,**l**)(-H,**l**)	

(30) Phonotactic constraint: Tone-coda co-occurrence restriction

PHONO$_A$: Highly ranked constraints which induce the tone-coda co-occurrence restriction. Here, I adopt Agbayani (1997)'s analysis which posits a licensing condition on [stiff]. He uses the following two undominated constraints based on Itô, Mester, and Padgett (1995):

a. **LICENSE [stiff]**: The feature [stiff vocal cords] ([stiff]) must be licensed.

b. **Obs]$_\sigma$ ⊃ [stiff]**: The specification [Obstruent] in coda ⇒ [stiff].

Note that since these constraints are undominated, inputs that violate this constraint simply do not have a phonetic realization. I suggest that the unacceptability of forms such as the ones in 0 is due to blocking effects. Detailed discussion and analysis of these facts are beyond the scope of this paper and are pending future research.

(31) Examples of avoidance of identical structure

a. lê la lê la

b. *lê la lê la

For the case of iec-Red, the following constraint is ranked above the other alignment constraints and accounts for two facts: 1) Even though the output is a disyllabic form, it is two separate words and thus exhibits the stress patterns and separability properties characteristic of two separate words, and 2) RED appears on the right side of BASE.

ALIGN-R (RED$_{iec}$, ϕ): For every RED$_{iec}$, there is a phonological phrase ϕ such that the right edge of RED$_{iec}$ coincides with the right edge of ϕ. This constraint ranks above the other alignment constraints. This accounts for RED$_{iec}$ appearing on the right side of BASE and the fact that RED$_{iec}$ and BASE are separate prosodic words and therefore have separate stress.

RED$_{iec}$ exhibits a number of properties which supports its status as a word (it can be separated from the base like most compounds can in Vietnamese; it receives its own stress, etc.).[17] For more detailed discussion of the syntactic separability of compounds in Vietnamese and the status of RED$_{iec}$, see Vu (1998).

[17] One possibility is that RED$_{i\%c}$ is simply a phonetically underspecified word which becomes specified in the context of another word. In cases where it is separated from the base, one possibility is to say that it raises covertly at LF to check features with the base and in doing so is filled in with phonetic material via correspondence relations.

The constraints ALIGN, IO-FAITH, PHONO_A, and FT-BIN (σ) are undominated except for iec-Red in which case ALIGN-R (RED_{iec}, ϕ) out-ranks these constraints.[18] Note that FT-BIN is always violated in iéc-Red for cases where the base is monosyllabic since the output is dysllabic but its prosodic structure is that of two separate words, each with one syllable.

Summary of highly ranked constraints:
Align-R (red_{iec}, ϕ) >> {Align, Ft-Bin, Phono_a ,IO-Faith} >> {relatively-ranked constraints}

4.2 *Relatively-ranked (conflicting) constraints*
Following the work of Itô and Mester (1995), I employ the concept of a cophonology to model what appears to be conflicting rankings of constraints within a single language. The reduplication processes examined belong to one of two cophonologies. An example is the difference in the status of different kinds of borrowed words in a language with respect to a core-periphery distinction (native words and older borrowed words being considered 'core' while recently borrowed words are more at the periphery of a language's vocabulary) as discussed in Itô and Mester (1995).

(32) Relatively ranked constraint: Base-Reduplicant Identity
BR-IDENT: Corresponding elements in the base and reduplicant must match in segmental and tonal features (register and contour values).

(33) Relatively ranked constraints: Dispreference of marked structure
a. ***MARKED**: No marked features can be associated with segments. I use this as an abbreviation for the following set of constraints against marked features appearing on segments (Agbayani (1997)):
b. Tonal markedness constraint:
***[+stiff]**$_\mu$: No [+stiff] features can be associated with moraic segments. This predicts that the unmarked tonal contour is l.
c. Vocalic markedness constraints:
***V-PL/LAB, *V-PL/DOR**: No [labial, dorsal, coronal, radical] specification for V(owel)-Place, respectively.

I propose that each of the four reduplication processes in Vietnamese belongs to one of these two cophonologies[19]:

(34) Cophonolologies in reduplication
a. BR-Ident >> *Marked
b. *Marked >> BR-Ident

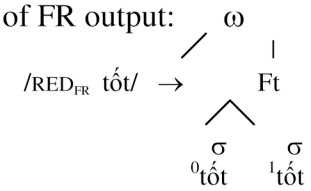
[18] Another formulation of these facts is to say that the ranking: ALIGN-R (RED_{iec}, ϕ) >> ALIGN, IO-FAITH, PHONO_A, FT-BIN (σ) >> {relatively ranked constraints} holds for all four processes.
[19] Note that in OT terms, different cophonologies can simply be different rankings of a common set of constraints.

4.3 *Example tables*

Below, I show which of the two cophonologies each process belongs to by computing example tables on data presented in previous sections. For the sake of expositional clarity, the bulk of the argument will be only for which cophonology the process belongs to and hence will not include every constraint mentioned in the previous section. In these tables, only the candidates relevant to arguing the ordering of BR-Ident and *Marked are considered. These constraints determine the segmental and tonal content of RED but does not bear on the prosodic structure of RED in the output.

Full reduplication: In all of these cases, the output appears to be a word with a single disyllabic foot. I refer to the reduplicant in FR as RED_{FR}. In all of these cases, since the output is a single prosodic word there can be no syntactic material inserted in between BASE and RED since that would violate the Prosodic Hierarchy by having a word dominated by a foot. The stress pattern 01 for these forms follows from the prosodic structure since there is only one stress assigned to each foot. The prosodic structure of the output is given in 0. I propose that FR is in the cophonology in which base-reduplicant identity is preferred over not having marked structure.

(35) Prosodic structure of FR output:



(36) Example: FR output form ⁰ít ¹ít= ‘small in amount’ (+att)

*MARKED Constraints (unranked)

	/RED_{FR}+ ít/ (+H,**h**)	BR-IDENT	*[+stiff]$_{\mu}$	*V-PL/LAB	*V-PL/DOR
a.	in ít (+H,**l**)(+H,**h**)	*!* (n, **l**)	* (**h**)	✓	✓
b.	Ît ít (-H,**h**)(+H,**h**)	*! (-H)	** (**h, h**)	✓	✓
c. →	ít ít (+H,**h**)(+H,**h**)	✓	** (**h, h**)	✓	✓

Candidates a) and b) in (36) are ruled out because they all have at least one violation of BR-Ident either because of differences in segmental or tonal material. The tension between identity and the dispreference of marked material is manifested in the fact that in order to fully satisfy BR-IDENT, there has to be at least two violations of *MARKED constraints, namely the **h** in the tones. This example shows that BR-IDENT is ranked above *MARKED. Candidates with a vowel change in the reduplicant are not considered since any such change will result in a violation of BR-IDENT whether they result in a violation of *MARKED or not.

(37) Example: FR output form ^{0}tốt ^{1}tốt = ‘good, useful’ (+att)

Undominated constraints (unranked)

	/RED$_{FR}$+tốt/ (+H,**h**)	IO-FAITH	PHONO$_A$	FT-BIN	ALIGN	BR-IDENT	*MARKED
a.	ω | Ft 2 σ σ ^{0}tôn ^{1}tôn	*!* (n,tone 1)	✓	✓	✓	✓	**** (o,o)
b.	ω | Ft 2 σ σ ^{0}tôt ^{1}tốt	✓	*! (tone 1)	✓	✓	* (tone 1)	***** (o,o, tone 2)
c.	ω 2 Ft Ft | | σ σ ^{1}tốt ^{1}tốt	✓	✓	*!	✓	✓	****** (o,o, tone 2, tone 2)
d.	ω | Ft 2 σ σ ^{1}tốt ^{0}tốt	✓	✓	✓	*! (Align-L (RED, ω))	✓	****** (o,o, tone 2, tone 2)
e.	ω | Ft 2 σ σ ^{0}tắt ^{1}tốt	✓	✓	✓	✓	*! (V)	**** (o, tone 2, tone 2)
f. →	ω | Ft 2 σ σ ^{0}tốt ^{1}tốt	✓	✓	✓	✓	✓	****** (o,o, tone 2, tone 2)

Candidates a) – d) are ruled out by the undominated constraints. e) (here, V = i, e, ɛ, a, i.e., the vowels with none of the marked V-Place features) is ruled out by a fatal violation of BR-IDENT.

Reduplication due to Emergence of Unmarked effects: Agbayani (1997) attributes base-reduplicant mismatch phenomena in Vietnamese reduplication to the interaction of TETU effects (McCarthy and Prince (1994b)) with the Normal Application of Phonological Constraints scheme (McCarthy and Prince (1995)). Specifically, TETU effects refer to the important empirical observation that unmarked effects which would otherwise be invisible in most parts of a grammar arise in unexpected ways in certain domains of a language

(e.g., reduplication). In correspondence-theoretic terms, these effects are accounted for as unmarked constraints out-ranking base-reduplicant identity constraints as shown in (38).

(38) The Emergence of the Unmarked ranking schema:
I-O FAITHFULNESS >> PHONO-CONSTRAINT$_A$ >> B-R IDENTITY

In the above schema, Phono-Constraint$_A$ are the hidden phonological constraints on unmarkedness. In the case of Vietnamese, for tones, the feature [+stiff] is more marked then [-stiff]; for vowels and codas, [+labial] and [+dorsal] are more marked then [+coronal] and [+radical]. Therefore, the restriction against the marked features may force violations of base-reduplicant identity. This notion can also be used to predict the various vowel and coda replacement in a number of reduplication processes. The result of this mismatch then is partial reduplication of BASE.

This scheme is itself dominated by even more highly ranked constraints on syllable structure, the segment-tone co-occurrence restriction mentioned above, and a faithfulness constraint of BASE in the output to BASE in the input. In my analysis, I show that these constraints are in fact undominated for all reduplication process in Vietnamese. This scheme (given below) is well-attested and referred to in the literature as the Normal application of Phonological Constraints.[20] (McCarthy and Prince (1995)) It is characterized as the following ranking schema:

(39) Normal Application of Phonological Constraints ranking schema:
PHONO-CONSTRAINT$_B$ >> I-O FAITHFULNESS >> B-R IDENTITY

The effect of this schema is that it induces similar effects as in (39), in that it contributes to the mismatch of the base and stem, but only to the extent that the mismatch does not violate a number of highly-ranked phonological constraints which are active throughout the grammar. The schemata (38) and (39) combine to produce the ranking (40). See Agbayani (1997) for the details of this analysis.

(40) Interaction of the two schemata:
PHONO-CONSTRAINT$_B$ >> I-O FAITHFULNESS >> PHONO-CONSTRAINT$_A$ >> B-R IDENTITY

Although Agbayani (1997)'s analysis correctly predicts the form of some reduplications in Vietnamese, its main limitation is that the data set itself is only a small subset of the overall system of reduplication in Vietnamese. In fact, there are examples in his data set which actually belong to other reduplication processes which are overlooked in his analysis. For example, the form *đẹp đẹp* is ruled out in his analysis but is a possible form in Vietnamese. The point is that although this form is not produced by TETU, it is produced by FR. The main difference is in register in which the forms are used.

Furthermore, his data set is restricted not only to single morpheme reduplications, but only one *type* of such reduplication; the other types such as FR and iec-Red are not accounted for. Other parameters such as stress and syntactic separability are also not

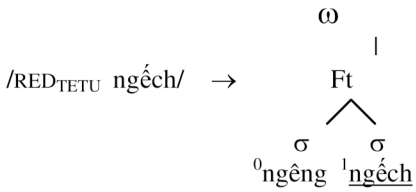
[20] As opposed to the Underapplication or Overapplication of Phonological constraints in which case faithfulness constraints are ranked below identity constraints or above normally highly-ranked constraints, respectively. (McCarthy and Prince (1995))

analyzed. Below, I will show that by a simple extension of some of his arguments and by reconsidering some of the constraints he posits in a more general context, I can account for additional processes as well as parameters along which these processes vary. In doing this, problematic cases such as the large number of cases in which marked structure *does* appear in RED can be explained.

I propose that TETU is prosodically very similar to FR in that it exhibits the same stress patterns (01) and does not have separability effects. However, it crucially differs from FR in that the *MARKED constraints are ranked above BR-IDENT. Adopting Agbayani's analysis, this forces base-reduplicant mismatch. The mismatch must be minimal in order for the candidate to be successful, i.e., it must maximally satisfy BR-IDENT. I call the reduplicant RED_{TETU} to distinguish the two different types of reduplicants.

Like FR, TETU reduplication is occurring at the syllabic level rather then the foot or prosodic word level. This gives TETU the same properties as FR with respect to the stress patterns seen in the output and its syntactic inseparability.

(41) Prosodic structures involved in TETU:



I propose that FR is in the cophonology in (42b) in which BR-IDENT is outranked by *MARKED.

(42) Example: TETU output form ^{0}thăng ^{1}thẳng= 'straight' (+att)
Input: /RED_{TETU}+ thẳng / (tone is (+H**,lh**))

*Marked Constraints (unranked)

	/RED_{TETU}+ th£ng/ (+H**,lh**)	*[+stiff]$_{\mu}$	*V-PL/LAB	*V-PL/DOR	BR-IDENT
a.	^{0}thăng ^{1}thẳng (-H**,lh**)(-H**,lh**)	**!	✓	✓	✓
b. →	^{0}thăng ^{1}thẳng (-H**,l**) (-H**,lh**)	*	✓	✓	*!

Candidate a) in (42) has two violations of *MARKED constraints which proves to be fatal. b) on the other hand, only has one violation of *MARKED and is the winning candidate even though it has one more BR-IDENT violation. Again, candidates with vowel changes in RED are not considered because any such change would result in additional violations of BR-IDENT and possibly *MARKED as well. This shows that the *MARKED constraints are indeed ranked above BR-IDENT in this process.

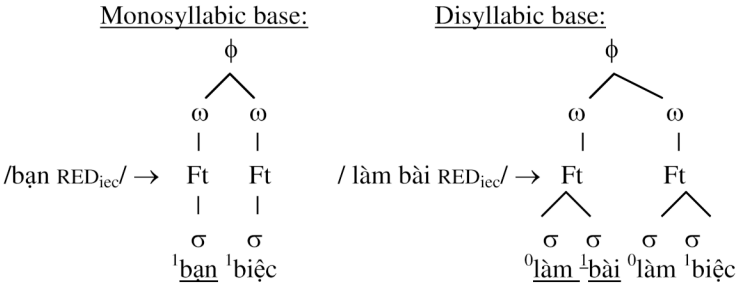
A table similar to (37) can be computed for ^{0}*ngêng* ^{1}*ngếch*, except in this case, *MARKED would be reanked above BR-IDENT

Reduplication via Melodic Overwriting I: iec-reduplication: iec-Red is probably the most productive of all the reduplication patterns in Vietnamese as it can regularly apply to any lexical word. I show below that the process belongs to the cophonology in which

*MARKED constraints are outranked by BR-IDENT. This ranking can be argued for using cases where the base has two syllables.

As mentioned above, this process involves the placement of an additional prosodic word (RED_{iec}) rather than the appending of phonological material within a prosodic word. Thus, continuing to assume that each foot has one stress assigned, the stress pattern 0101 for disyllabic bases follows from the prosodic structure of the reduplicative form. The syntactic separability of the form is also explained by the fact that it is possible to insert a word in between word (within a phonological phrase ϕ) without violating the integrity of the Prosodic Hierarchy. Note that this structure also accurately predicts the inseparability of the morphemes in *làm bài* and in *làm biệc* since each of these pairs of syllables is within the same prosodic word. Similar accurate predictions follow from prosodic structure for cases with monosyllabic case as well.

(43) Prosodic structure of iec-Red outputs



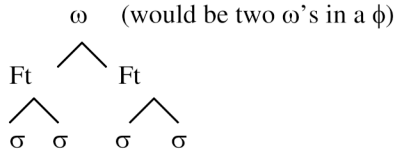
(44) Example: i%ɔc-Red output form *^{0}bún ^{1}chả ^{0}bún ^{1}chiếc* = 'noodle and grilled pork dish' (+pej)

*MARKED Constraints (unranked)

	/ bún chả + RED_{iec}/	BR-IDENT	*[+stiff]$_{\mu}$	*V-PL/LAB	*V-PL/DOR
a.	^{0}bún ^{1}chả ^{0}bun ^{1}chiếc	****!(iɤ, k,tone1,2)	** (tone2 x2)	** (u,u)	*** (u,u,iɤ)
b. →	^{0}bún ^{1}chả ^{0}bún ^{1}chiếc	***(iɤ,c,tone2)	*** (tone 2 x3)	** (u,u)	*** (u,u,iɤ)

Candidate a) has four violations of BR-IDENT which is fatal. Candidate b) has one more violation of *MARKED, but it has one fewer violation of BR-Ident and thus is the winning candidate.

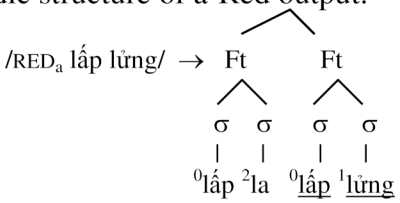
For space considerations a full table for iec-Red will not be presented. The ranking would be: ALIGN-R (RED_{iec},ϕ) >> {IO-FAITH, $PHONO_{A}$, FT-BIN, ALIGN} >>BR-IDENT>> *MARKED. The main difference between such a table and a table like 0 would be that the first alignment constraint would prevent structures such as:



Reduplication via Melodic Overwriting II: a-reduplication: Although a-Red is productive in the sense that given a form, it is predictable whether this process can operate on the form or not, it is limited to disyllabic bases that are simple verbs. a-Red is in the same cophonology as iec-Red (and hence FR as well).

As was the case above, the stress pattern 0201 and lack of syntactic separability of a-Red forms follow directly from the prosodic structure I propose for the reduplicative forms (given in (45)). Since each foot is assigned one stress, we have the alternation oxoy where x,y ≠ 0. That there is a secondary/primary stress distinction, and that there cannot be a word inserted in between BASE and RED, give evidence to the hypothesis that copying is indeed happening at the foot level and that the entire utterance is dominated by one prosodic word. That BASE and RED cannot be separated by another word is because a word cannot be inserted inside another word.

(45) Prosodic structure of a-Red output: ω



(46) Example table for a-Red output form *^{0}lấp ^{2}la ^{0}lấp ^{1}lửng* = 'to bob up and down/equivocate'(+emph)[21] *MARKED Constraints (unranked)

	/RED$_A$ + lấp lửng/	BR-IDENT	*[+stiff]$_\mu$	*V-PL/LAB	*V-PL/DOR
a.	^{0}lấp ^{2}là ^{0}lấp ^{1}lửng	***! (a,a:,tone1)	*** (tone 2x3)	** (o,o)	** (o,o)
b. →	^{0}lấp ^{2}la ^{0}lấp ^{1}lửng	** (a:, tone 2)	*** (tone 2 x3)	*** (o,o,o)	*** (o,o,o)

In (47), candidate a) has three violations of BR-IDENT, which proves to be fatal since the winning candidate only has two violations of this constraint even though a) has two fewer *MARKED violations. This shows that BR-IDENT is ranked above the *MARKED constraints. The table below illustrates the prosodic structural constraints for the same input.

[21] Thanks to Cao Xuân Hạo for pointing out this example.

(47) Example: a-Red output form ^{0}*lấp* ^{2}*la* ^{0}*lấp* ^{1}*lửng* = 'to bob up and down, to equivocate' (+emph) (full table)

	/RED$_{A}$ + *lấp lửng*/	IO-FAITH	PHONO$_{A}$	FT-BIN	ALIGN	BR-IDENT	*MARKED
a.	ω ei Ft Ft 2 2 σ σ σ σ ^{0}lấp ^{2}la ^{0}lấp ^{1}la	**! (a:, tone 1)	✓	✓	✓	✓	**** (ɤ,ɤ, tone 2, tone 2)
b.	ω ei Ft Ft 2 2 σ σ σ σ ^{0}lấp ^{2}la ^{0}lấp ^{1}lửng	✓	*! (tone 1)	✓	✓	** (tone 1, tone 1, a:)	*** (ɤ,ɤ, tone 2, tone 3)
c.	ω ω ω | | Ft Ft Ft | | σ σ σ σ ^{1}lấp ^{1}la ^{0}lấp ^{1}lửng	✓	✓	*!	**! ALIGN-R (ω, Stem)	** (tone 1, a:)	***** (ɤ,ɤ, tone 2, tone 2, tone 3)
d.	ω ω | Ft Ft 2 2 σ σ σ σ ^{0}lấp ^{1}la ^{0}lấp ^{1}lửng	✓	✓	✓	*! ALIGN-R (ω, Stem)	** (a:, tone 1)	***** (ɤ,ɤ, tone 2, tone 2, tone 3)
e.	ω ei Ft Ft 2 2 σ σ σ σ ^{0}lấp ^{2}lá ^{0}lấp ^{1}lửng	✓	✓	✓	✓	***! (a,a:, tone 1)	**** (ɤ, tone 2, tone 2, tone 3)
f. →	ω ei Ft Ft 2 2 σ σ σ σ ^{0}lấp ^{2}lá ^{0}lấp ^{1}lửng	✓	✓	✓	✓	** (a:, tone1)	***** (ɤ,ɤ, tone 2, tone 2, tone 3)

In (47) a) – d) show that violation of any of the first four constraints is fatal; therefore, whether they are ranked relative to one another or not (I give no argument), they are indeed undominated. e) and f) show that BR-IDENT dominates *MARKED and hence, as much of the segmental as well as tonal material must be copied while preserving the prespecified shape of RED, in this case the fixed segment *a:* in the second syllable in RED.

5 Discussion and summary

In the above analysis, I have shown that by examining the prosodic structure of the reduplicative forms in Vietnamese, a number of seemingly disparate properties of these forms can be accounted for. What is particularly important is that copying at different levels in the prosodic structure can result in the output form obtaining different characteristics. For example, copying at the prosodic word level results in the syntactic separability of BASE and RED, whereas copying at any lower level will not. That reference to prosodic structure can give such a unified account supports the Prosodic Hierarchy Hypothesis. Moreover, this analysis brings to light the general importance of such structure in the theory of reduplication cross-linguistically.

I have also shown that FR, a-Red, and iec-Red belong to the cophonology in which constraints against markedness are dominated by constraints forcing an identity relation between BASE and RED. TETU is shown to be in the cophonology in which (as the name of the process implies) constraints against markedness emerge in the output forms. One of the implications of this find is the idea that it may be possible to have conflicting rankings of the same constraints within the same language. This has been shown to be possible in a somewhat different context in Itô and Mester (1995), in which they show that Faithfulness constraints may vary in its ranking depending on how well assimilated a lexical item is in a language's vocabulary. In this case, it is a matter of constraints against markedness arising in certain reduplication patterns and not in others.

6 Conclusions and questions

That there may be rerankings of constraints within a single language raises a number of questions for the architecture of OT as an explanatory system. For example, does allowing conflicting constraint rankings make the system too powerful? Although such an alteration to the theory may not be conceptually desirable, what alternatives are there given the assumptions made by OT? With these questions in mind, I suggest that a more productive line of research would be to explore which parts of the phonological system (and the grammar in general) are optimality theoretic and which are not.

One of the issues the data raises, but which was not discussed is that the Melodic Overwriting (MO) processes (a-Red and iec-Red) are blocked from operating on certain special forms which would normally be predicted to be valid inputs. I am referring specifically to the starred examples listed in 0 and 0, repeated here in (48).

Blocked forms

a.	liếc	(+H,**h**)	to scam
b.	*[1]<u>liếc</u> [1] liếc	(+H,**h**)(+H,**h**)	
c.	tiếc	(+H,**h**)	to regret
d.	*[1]<u>tiếc</u> [1] tiếc	(+H,**h**)(+H,**h**)	
e.	[1] tiêng [1]tiếc	(+H,**l**)(+H,**h**)	to regret (+pej)
f.	leː laː	(+H,**l**)(+H,**l**)	to wander drunkenly
g.	*leː laː <u>leː laː</u>	(+H,**l**)(+H,**l**)(+H,**l**)(+H,**l**)	

That there is no alternative form and that the meanings (e.g., 'to scam' (+pej)) would have to be expressed paraphrastically or by the use of an alternative but

unpredictable lexical form suggests a promising topic for future study. It is not clear how such blocking phenomena such as these can be explained within the OT framework since there is no way for the system to 'crash,' i.e., to have the winning candidate to be ∅, the empty set. Although Prince and McCarthy (1993) have alluded to a solution employing the notion of a Null Parse, this approach seems unsatisfactory since the result is still an empty category (e). However, e is still something that is being generated (in this case by the phonological system), it just happens to not have any phonetic realization. The output produced by such a process is not the same as the empty set since it is very possible for an empty element to have syntactic and semantic features which would have effects in other areas of the grammar.

In these instances of blocking, it does not seem that there is an empty element being generated. First, there is no syntactic reflex that is caused by such an empty form. Second, there is often an unpredictable replacement form such as in example (69e). Although I do not provide an explanation for these observations, the fact that this sort of blocking in MO processes occurs in other languages such as Turkish (Meltem Kelepir, p.c.) warrants further investigation. Note that this sort of blocking is distinct from what happens in languages like Abkhaz (Bruening (1996)) in which the MO processes that are blocked have a regular alternative process which handles the blocked cases. Therefore, as with the problem of conflicting constraints mentioned above, this need to 'crash' rather than to just generate an empty element seems to be a problem that is irreconcilable within the OT system.

Finally, another point of interest is the observation that iec-Red is allowed on any lexical item except simple disyllabic verbs, in which case, a more specific rule, namely a-Red, is applied. Although this sort of pattern has been well attested and is explained by notions like the Elsewhere Condition, what is of interest is that a-Red and iec-Red do not produce the same changes in meaning on a given input. The rule a-Red gives an emphatic meaning to BASE whereas iec-Red produces a pejorative meaning.

References

Agbayani, Brian. 1997. Phonological Constraints and Base-Reduplicant Non-Identity: The Case of Vietnamese. Ms. University of California, Irvine.

Alderete, John, etal. 1997. Reduplication and fixed segmentism. Ms., University of Massachusetts, Amherst.

Alves, Mark. 1997b. Tonal features and the development of Vietnamese tones. Ms. University of Hawaii.

Bruening, Benjamin. 1996. Abkhaz Mabkhaz: m-reduplication in Abkhaz, Weightless Syllables, and Base-Reduplicant Correspondence. In Benjamin Bruening,Yoonjung Kang, and Martha McGinnis, eds., *MIT Working Papers in Linguistic:30*. 291-329.

Burton, Strang. 1992. Reduplication and the representation of tone in Vietnamese. Ms., Brandeis University.

Cao Xuân Hạo. 1997. Trọng âm và các quan hệ ngữ pháp trong Tiếng Việt. (Stress and its relation to grammar in Vietnamese) Ms., Saigon National University.

Đoàn Thiện Thuật. 1980. *Ngữ âm Tiếng Việt*. [Vietnamese Phonetics] Hanoi: Nhà Xuất Bản Đại Học và Trung Học Chuyên Nghiệp.

Edmondson, Jerold and Nguyễn Văn Lợi. 1997. Tones and voice quality in modern northern Vietnamese: instrumental case studies." *Mon-Khmer Studies 28.*

Halle, Morris and K. N. Stevens. 1971. A note on laryngeal features. *Quarterly progress report No. 101*. 198-213. Research Laboratory of Electronics, MIT.

Han, Meiko S. 1966. Vietnamese vowels. In *Studies in the phonology of Asian languages* technical report 4. Done at Acoustics Phonetics Research Laboratory, University of Southern California.

Han, Meiko S. 1969. Vietnamese tones. In *Studies in the phonology of Asian languages* technical report 8.

Haudricourt, André. 1954. De l'origine des tons en Viêtnamien. *Journal Asiatique 242*: 69-82.

Hayes, Bruce. 1989. The prosodic hierarchy meter. In *Phonetics and Phonology: Rhythm and Meter*, Paul Kiparsky and Gilbert Youmans, eds., 1:201-260. New York: Academic Press.

Hoàng Văn Hành. 1985. *Từ láy trong Tiếng Việt*. [Reduplicative forms in Vietnamese] Hanoi: Nhà Xuất Bản Khoa Học Xã Hội.

Hoàng Văn Hành, et. al. 1995. *Từ điển từ láy Tiếng Việt*. [Dictionary of Reduplicative Forms]. Hanoi, Vietnam.

Itô, J., A. Mester, and J. Padgett. 1995. Licensing and underspecification in Optimality Theory. *Linguistic Inquiry 26*:571-613.

Itô, J. and A. Mester. 1995. The core-periphery structure of the lexicon and constraints on reranking. In Beckman, Dickey, and Urbancyck, eds., *University of Massachusetts Optimality Papers 18*. Amherst, MA.

McCarthy, John and Alan Prince. 1993b. Generalized alignment. *Yearbook of Morphology*, pp. 79-153.

McCarthy, John and Alan Prince. 1994a. An overview of prosodic morphology. Parts 1 and 2: Template form in reduplication; Template satisfaction. Talks at Utrecht Prosodic Morphology Workshop.

McCarthy, John and Alan Prince. 1994b. The emergence of the unmarked: Optimality in prosodic morphology. In Merce Gonzalez ed., *Proceedings of the North-East Linguistics Society 24*. Amherst, MA.

McCarthy, John and Alan Prince. 1995. Faithfulness and reduplicative identity. In UMOP 18.

Nespor, Marina and Irene Vogel. 1982. Prosodic domains of external sandhi rules. In *The Structure of Phonological Representations 1*, Harry van der Hulst and Norval Smith, eds. Dordrecht: Foris.

Ngô Thanh Nhàn. 1984. The Syllabeme and patterns of word formation in Vietnamese. Ph.D. dissertation, New York University.

Nguyễn Đinh Hoà. 1980. Language in Vietnamese Society. *Vietnam Culture Series I*, Asia Books, 63-108.

Nguyễn Kim Thản. 1997. *Nghiên cứu ngữ pháp Tiếng Việt*. [Studies in Vietnamese Grammar] Hanoi: Nhà Xuất Bản Giáo Dục.

Nguyễn Tài Cẩn. 1996. *Ngữ pháp Tieng Việt*. [Vietnamese Grammar] Hanoi: Nhà Xuất Bản Đại Học Quốc Gia Hà Nội.

Noyer, Rolf. 1995. Vietnamese 'morphology' and the definition of Word. Ms., Brown University.

Phạm Hoa. 1997. The coronal-velar relationship in Vietnamese dialects. M.A. Thesis, University of Toronto.

Prince, Alan and Paul Smolensky. 1993. Optimality Theory: constraint interaction in generative grammar. Ms. Rutgers University and the University of Colorado, Boulder.

Selkirk, Elisabeth. 1978. On prosodic structure and its relation to syntactic structure. In *Nord Prosody II*, T. Fretheim, ed., 111-140. Trondheim, Norway: TAPIR.

Thompson, Laurence C. 1965. *A Vietnamese grammar*. Seattle: University of Washington Press.

Trần Mai Hương Aurelie. 1967. Stress, tones, and intonation in South Vietnamese. Ph.D. Dissertation, Australian National University.

Urbanczyk, Suzanne. 1997. Morphological templates in reduplication. In K. Kusumoto, eds., *North-Eastern Linguistics Society 27*: 369-382.

Võ Xuân Hào. 1997. Nghiên cứu về chức năng của thanh điệu Tiếng Việt. [Studies in the function of tones in Vietnamese]. Ph.D. Dissertation, Hanoi National Unversity.

Vũ Thế Thạch. 1994. Consonant copying and tone harmony in Vietnamese reduplicatives. In Martha Ratliff and Wric Schiller, eds., *Papers from the First Annual Meeting of the Southeast Asian Linguistics Society.* Tempe: Arizona State University: 381-395.

Vu, Sonny Xuan. 1998. Doing splits at PF: Compound separability in Vietnamese and Pacoh. Ms. MIT.

Wannamacher, Mark. 1997. Aspects of tone and voice quality: an autosegmental account. Arlington: Texas. University of Texas, Arlington MA thesis.

Yip, Moira. 1992. Reduplication with fixed melodic material. *Proceedings of the North Eastern Linguistics Society 22*: 459-476.