

## LOANWORDS AS PSEUDO-COMPOUNDS IN MALAGASY\*

ANDREW MARTIN

University of California, Los Angeles  
amartin@humnet.ucla.edu

Reduplication in Malagasy copies a two-mora-long section of the base (e.g., alik<sup>1</sup>a ‘dog’ → alik<sup>1</sup>al<sup>1</sup>ik<sup>1</sup>a ‘something like a dog’). Although the reduplicant is usually left-anchored to the main-stress syllable, there are two types of words in the language that place the reduplicant elsewhere (anchored to secondary stress): finally-stressed loanwords, and native compounds. I argue that loanwords behave like compounds in reduplication because they are in fact analyzed by speakers as compounds. This reanalysis is triggered by the fact that the stress pattern in these loanwords is otherwise found only in compounds.

### 1. INTRODUCTION

When speakers incorporate words from another language into their own language, they are often confronted with structures that are illegal in the recipient language—for example, new phonemes, banned consonant clusters, or abnormal stress patterns. Depending on the language, speakers resort to a variety of strategies to solve these problems, ranging from simply allowing the illegal structure, to changing it into a legal one, to deleting it outright. Another, less common, solution is to reanalyze the foreign word, positing morphological structure that would justify the illicit structure.

I will argue that just such a situation holds in some words borrowed into Malagasy. Words borrowed from French with stress on the final vowel, a pattern that in the native vocabulary only occurs in compounds, are treated as compounds on the basis of this surface resemblance, thus making them *pseudo-compounds*. Evidence for this comes from the fact that finally-stressed loanwords behave like compounds when they are reduplicated—only in compounds and loanwords does the reduplicant fail to copy the main-stress syllable.

In the next section, I begin by describing the stress system and the way that native words (other than compounds) reduplicate in Malagasy. In section 3, I show that some loanwords, those with an anomalous stress pattern, reduplicate differently than native words. In section 4, I demonstrate that the behavior of native compounds in Malagasy is different than that of monoroot words, but

---

\* Many thanks to my consultant Noro Ramahatafandry, for her tireless reduplication and many insights into her language and culture. All of the Malagasy data in this paper is drawn from my elicitation sessions with her. I am also grateful to Kie Zuraw for much useful discussion, and to the audience at AFLA XII for ideas and comments. Any remaining errors are my own.

strikingly similar to the behavior of the irregular loanwords. In section 5, I argue that this similarity is the result of pseudo-compounding, and discuss the theoretical ramifications of this proposal. Finally, section 6 consists of a summary of the findings presented here.

## 2. REDUPLICATION IN NATIVE WORDS

Malagasy has a single type of reduplication, which is unusual in that it freely applies to verbs, nouns, and adjectives. The meaning of reduplicated words can be idiosyncratic, especially for common words, but the patterns listed in (1) are most common.

### (1) Reduplication semantics

<i>verb</i>	<i>noun</i>	<i>adjective</i>
‘sort of Xing’	‘something that acts like X’	‘somewhat X’
‘Xing but not seriously’	‘something resembling X’	‘very X’
‘Xing intermittently’		

Phonologically, reduplication appears to behave identically regardless of the category of the base; I will therefore analyze the process as a single phenomenon.

Because stress plays a crucial role in reduplication, and in the argument developed in this paper, a few words on the Malagasy stress system are necessary. In native Malagasy monoroot words, stress is entirely predictable—main stress falls on the penultimate mora, as the examples in (2) illustrate.

### (2) Stress on penultimate mora<sup>1</sup>

- (a) mànadála ‘to fool’
- (b) alík<sup>l</sup>a ‘dog’
- (c) mamái ‘to punish’
- (d) milaláu ‘to play’

Although some words surface with antepenultimate stress (see (3) below), I follow Erwin (1996) in assuming that these words are consonant-final, and that the final vowels are epenthesized because of Malagasy’s ban on codas.

### (3) Antepenultimate stress

- (a) námana ‘friend’
- (b) ma<sup>m</sup>p<sup>l</sup>ánatra ‘to teach’
- (c) sífaka ‘lemur’

<sup>1</sup> Malagasy examples are given in broad IPA (voiceless vowels are not marked as such), *not* native orthography. Stress is marked with accents. I assume that NC clusters are actually <sup>NC</sup> prenasalized stops.

Stress is thus assigned as follows: moraic trochees are constructed from right to left, with syllables containing epenthetic vowels being extrametrical. Primary stress is assigned to the rightmost foot. The main stress, then, will always end up on the penultimate (non-epenthetic) mora. The only exceptions to this generalization in native monoroot words are words that contain only a single short vowel (e.g., [fú] ‘heart’), where the final vowel is the only vowel, and so must be stressed, presumably due to a requirement that all words bear stress.<sup>2</sup>

In native, monoroot words, reduplication copies the main-stress syllable and the next syllable, unless that syllable contains an epenthetic vowel. Examples are given in (4).<sup>3</sup>

#### (4) Monoroot reduplication

	<u>simple</u>	<u>reduplicated</u>	
(a)	mànadála	mànadàladála	‘to fool’
(b)	alík <sup>h</sup> a	alík <sup>h</sup> alík <sup>h</sup> a	‘dog’
(c)	mamáí	mamàifái <sup>4</sup>	‘to punish’
(d)	mìlaláu	mìlalàwláw	‘to play’
(e)	námana	nàmanámana	‘friend’
(f)	ma <sup>m</sup> p <sup>h</sup> ánat <sup>r</sup> a	ma <sup>m</sup> p <sup>h</sup> ánap <sup>h</sup> ánat <sup>r</sup> a	‘to teach’
(g)	sífaka	sífatsífaka	‘lemur’

In these cases, the reduplicant always copies the main-stress foot, which means that it always contains at least two moras (unless the base itself has only one mora, as in [fú] ‘heart,’ which reduplicates as [fùfú]). By itself, this fact does not entail that there be a size minimum on the reduplicant; if reduplication simply copies material from the main-stress syllable to the end of the word (modulo epenthetic vowels), the two-mora size minimum would emerge as a consequence of the fact that stress always falls on the penultimate mora. In the next section, however, I present evidence from loanwords that there is in fact a size restriction on reduplicants.

### 3. REDUPLICATION IN LOANWORDS

The French colonization of Madagascar, beginning in the nineteenth century, resulted in a heavy influx of French loanwords which has continued to the present day. Most French words have been borrowed with final stress, even where this would result in an illegal stress pattern in Malagasy: [sòkolá] ‘chocolate’ (from French *chocolat*) and [zàvuká] ‘avocado’ (from French *les avocats*), for example, bear primary stress on the final mora, a pattern which is never seen in native monoroot words.

<sup>2</sup> In certain inflected forms (such as imperatives), stress can shift to a final light syllable, but such forms can never serve as a base for reduplication—in effect, inflection applies after reduplication.

<sup>3</sup> The first copy in each reduplicated form is underlined for clarity; I intend no theoretical claim regarding which copy is reduplicant and which base, as it is not crucial to the issues discussed here.

<sup>4</sup> Underlyingly /maN + fai/, this verb surfaces as [mamai] due to nasal substitution; the initial consonant of the root surfaces faithfully only in the second copy of the reduplicated form.

In section 2 I showed that in native words, the reduplicant is left-anchored to the main-stress syllable. If this requirement holds of loanwords, we expect [sòkolàlá] as the reduplicated form of [sòkolá]; surprisingly, the grammatical form is [sòkosòkolá], in which reduplication targets the secondary stress, and the main-stress syllable is not copied at all. The examples in (5) show that this true in general of loanwords whose final mora is stressed.

(5) Reduplication in finally-stressed loanwords

	<u>simple</u>	<u>reduplicated</u>	
(a)	sòkolá	sòkosòkolá	‘chocolate’
(b)	zàvuká	zàvuzàvuká	‘avocado’
(c)	làkiré	làkilàkiré	‘chalk’
(d)	sàribó	sàrisàribó	‘coal’

It is not the case that *all* loanwords reduplicate in this way; those that have unmarked stress patterns (i.e., stress on the penultimate mora) reduplicate in the same way as native words:

(6) Reduplication in penultimate-stressed loanwords

	<u>simple</u>	<u>reduplicated</u>	
(a)	gilási	gilàsilási <sup>5</sup>	‘ice cream’
(b)	òpitáli	òpitàlitáli	‘hospital’
(c)	làkulúsi	làkulùsilúsi	‘bell’

Above I noted that in native monoroot words, the reduplicant always contains at least two moras. The fact that \*[sòkolàlá] is ungrammatical suggests that there is a two-mora minimum size requirement: in [sòkolá], the main stressed syllable [la], being monomoraic, is simply too small to be copied, and so the reduplicant targets the secondary stress instead. Such a size restriction is attested in other languages, although the specific repair varies in each case (see (7)-(9)).

(7) *Kinande*: make multiple copies of base (Mutaka and Hyman 1990, Mutaka 1994)

	<u>simple</u>	<u>reduplicated</u>	
(a)	eri-tuma	eri-tuma=tuma	‘to send’
(b)	eri-swa	eri-swaswa=swa	‘to grind’

<sup>5</sup> Unlike epenthetic vowels in native words, those in loanwords (like the final *-i* in *gilasi*) are often copied into the reduplicant (cf. native *nàmanámana*). It may be that epenthetic vowels in borrowed words have become lexicalized and now have a different status than vowels epenthesized in native consonant-final words.

(8) *Swahili*: copy non-root material (Park 1997)

	<u>simple</u>	<u>reduplicated</u>	
(a)	inuka	inuka- <u>inuka</u>	‘rise up’
(b)	m-binya	m-binya- <u>binya</u>	‘pinch him’
(c)	m-pa	m-pe- <u>mpe</u>	‘give him’

(9) *Siswati*: expand base through epenthesis (Kiyomi and Davis 1992)

	<u>stem</u>	<u>reduplicated</u>	
(a)	dlala	dlala-dlala	‘to play’
(b)	dla	dlayi-dla	‘to eat’

All of these languages require the reduplicant to be at least two syllables; I propose that Malagasy places a similar minimal size of two moras on the reduplicant. Because of the regularity of the native stress system, which ensures that there is always a base of adequate size, this requirement can only be seen through its effects on loanwords. This suggests that the first Malagasy speakers to encounter French words had already learned this generalization, even though the native vocabulary gives learners no relevant negative evidence—there are no native words with stress on the final mora, and therefore no alternations that would show a learner that a subminimal base may not be reduplicated. This may, then, be a case of the emergence of the unmarked,

The question I will address in the rest of this paper, however, does not concern the minimal size constraint that drives this unusual reduplication pattern, but rather the specific choice of repair that Malagasy speakers use to satisfy the constraint. As the examples in (7)-(9) demonstrate, there are a number of ways that a language could repair a subminimal base for reduplication. Malagasy could, for example, epenthesize a vowel in order to expand the base, but instead chooses to move the site of reduplication. But why did the first Malagasy speakers to reduplicate these words choose this particular repair? In the remainder of this paper, I will suggest that an accidental, surface resemblance between finally-stressed loanwords and native compounds has resulted in loanwords being reduplicated as compounds.

## 4. REDUPLICATION IN COMPOUNDS

Compounds in Malagasy typically put roots in head-modifier order. Examples follow in (10).

## (10) Compounds

- |     |   |                   |
|-----|---|-------------------|
| (a) | bà-kiráru<br>sock-shoe                    | ‘socks’           |
| (b) | tài-kí <sup>n</sup> tana<br>feces-star    | ‘meteor’          |
| (c) | vàva-fú<br>mouth-heart                    | ‘solar plexus’    |
| (d) | takèla-bí<br>sheet-metal                  | ‘sheet metal’     |
| (e) | màira-tsáina <sup>6</sup><br>lighted-mind | ‘intelligent’     |
| (f) | tsù-búlu<br>straight-hair                 | ‘straight-haired’ |
| (g) | manù <sup>m</sup> bi-kéna<br>cut-meat     | ‘to cut meat’     |
| (h) | mìtsu-d <sup>f</sup> ánu<br>blow-water    | ‘to bless’        |
| (i) | àru-dùà-búdi<br>basket-hole-bottom        | ‘wasteful person’ |

As these examples show, stress is assigned to each root as if it were in isolation; [vàva-fú] receives stress on the final mora because the root [fú] contains only a single vowel. This allows compounds to violate the generalization described in section 3; unlike monoroot words, compounds may place stress on the final mora. Thus, both compounds and loanwords may exhibit this anomalous stress pattern, albeit for different reasons: in compounds, the presence of morphological boundaries is responsible, while in loanwords it is the result of faithfulness to perceived stress in the donor language. Despite this difference in the motivating factors, however, I will argue that the surface resemblance between the two types of words leads speakers to treat them as the same for the purposes of reduplication.

When compounds are reduplicated, copying targets *only* the first member of the compound:

<sup>6</sup> The initial consonant of a non-initial root in a compound often undergoes fortition; thus, /saina/ ‘mind’ surfaces as [tsaina] when it is the second member of a compound. Other such fortitions include /h/ > [k] and /r/ > [d<sup>f</sup>].

## (11) Compound reduplication

	<u>simple</u>	<u>reduplicated</u>	
(a)	maira-tsáina	mairaira-tsáina (*maira-tsàinasáina)	‘intelligent’
(b)	tài-kí <sup>n</sup> tana	tàitài-kí <sup>n</sup> tana (*tài-kí <sup>n</sup> takí <sup>n</sup> tana)	‘meteor’
(c)	àru-dùa-búdi	àrukàru-dùa-búdi (*àru-dùadùa-búdi) (*àru-dùa-bú <b>di</b> búdi)	‘wasteful person’

Because of this requirement, the reduplicant never copies the main-stress mora in a compound, unlike monoroot words, in which the main-stress mora is *always* copied.

The simplest way to account for this would be to posit an ordering among morphological operations: first, reduplication applies to the head root, and then it combines with the other roots in the compound. This explanation is unsatisfactory, however, when the semantics of reduplication are taken into account. Even in cases where the second member of a compound is clearly the one being semantically modified by reduplication, it is still the first member that is the target of copying. For example, in the compound [tùe-d<sup>f</sup>átsi] ‘unpleasant personality,’ composed of the roots [tùe] ‘nature’ and [rátsi] ‘bad,’ the only grammatical reduplicated form is [tùetùe-d<sup>f</sup>átsi], in which the root meaning ‘nature’ is copied. Despite this, the reduplicated form can only mean ‘a somewhat unpleasant personality’—it is the meaning of the second root, meaning ‘bad,’ that is modified by reduplication, even though this root cannot be the target of copying (\*[tùe-d<sup>f</sup>átsirátsi] is ungrammatical). This suggests that the “target first root” restriction on reduplication is a phonological constraint on the surface form rather than a consequence of reduplication being ordered before compounding.<sup>7</sup>

Because the first root in a compound is targeted, a compound like [vàva-fú] ‘solar plexus’ reduplicates as [vávavàva-fú]. This is strikingly similar to loanwords like [sòkolá], which reduplicates as [sòkòsòkolá]. In the case of the compound, the reduplicant is anchored to a secondary stress as a result of the requirement that the initial, head element of the compound serve as the target of reduplication. The loanword, however, does not contain multiple roots and so the motivation for moving the site of reduplication (as opposed to repairing the subminimal base in some other way) is less clear. In the next section I will argue that although the loanwords are not truly compounds, they are in fact treated as compounds by Malagasy speakers.

## 5. PSEUDO-COMPOUNDS

<sup>7</sup> Another piece of evidence that reduplication does not apply to individual roots, which are then compounded: if the first member of the compound consists of a single light syllable, and the following syllable (the first syllable of the second member) is unstressed, the reduplicant copies material from both roots. For example, [bà-kiráru] ‘socks’ → [bàkibà-kiráru]. (If the syllable following the subminimal root is stressed, then only the first root is copied: [tsù-búlu] ‘straight-haired’ → [tsùtsù-búlu].)

The idea that the accidental phonotactic properties of a word may lead speakers to posit spurious morphological structure, even in the absence of any semantic justification, has been suggested several times in the literature (Prince 1980, Hayes 1995, Hammond 1999a, b, Zuraw 2002, Riggle and Munro 2004). A representative example is given by Hayes (1995), who points out that the word *gobbledy-gook* is perceived by most native English speakers as being composed of two “morphemes,” reflected in the use of a hyphen when the word is spelled, despite the fact that neither *gobbledy* nor *gook* is by itself a meaningful unit. This is due, he argues, to the fact that unstressed [i] is normally found only morpheme-finally; its presence within the word *gobbledy-gook* leads learners to assume that it must therefore be followed by a morpheme boundary.

In the case of Malagasy loanwords, something similar may have happened when words with anomalous stress patterns were initially borrowed; normally, stress on the final mora signifies a compound, and so French loans were interpreted as compounds for the purposes of reduplication. This may be the reflex of a learning strategy used by children; recent experiments (e.g., McQueen 1998) have shown that children learn phonotactic regularities early, and use this knowledge to parse speech into words (and, presumably, words into morphemes). If children learning Malagasy use stress to determine when a word is a compound, and access to this morphological parser persists into adulthood, it may explain why foreign words which seem to violate native stress rules are assumed to be compounds (or at least treated as if they were compounds).

It is less clear how this process of analogy is to be formalized. Have speakers simply noticed that all finally-stressed words in the language reduplicate in a special way, and extended this generalization to newly borrowed words, or are words like [sokola] actually stored as compounds (i.e., /soko+la/), and so reduplicate as compounds? At present the data is insufficient to distinguish between these possibilities, making this a fruitful area for future research.

## 6. CONCLUSION

I have shown that finally-stressed loanwords in Malagasy display irregular behavior when reduplicated. Although the fact that loanwords are irregular may be due to a minimal size requirement on the reduplicant, the precise way in which loanwords fulfill this requirement, I have argued, is due to a completely accidental resemblance to the stress patterns found in native compounds.

This suggests that people can learn accidental, surface-based generalizations and even productively extend them to novel forms, a phenomenon not easily accounted for by most current theories of phonology. The data presented here also shed light on a possible connection between the early stages of language learning and the treatment of foreign words by adult speakers.

## REFERENCES

- ERWIN, SEAN. 1996. Quantity and moras: An amicable separation. In M. Pearson and I. Paul (eds.), *The Structure of Malagasy, Volume I*. UCLA Working Papers.
- HAMMOND, MICHAEL. 1999a. English stress and cranberry morphs. Paper presented at LSA annual meeting, Los Angeles.
- . 1999b. *The Phonology of English: A Prosodic Optimality-Theoretic Approach*. Oxford: Oxford University Press.
- HAYES, BRUCE. 1995. *Metrical Stress Theory: Principles and Case Studies*. Chicago: University of Chicago Press.
- KIYOMI, S. and S. DAVIS. 1992. Verb reduplication in Swati. *African Languages and Cultures* 5. 113-124.
- MCQUEEN, JAMES. 1998. Segmentation of Continuous Speech Using Phonotactics. *Journal of Memory and Language* 39, 21-46.
- MUTAKA, NGESSIMO and LARRY HYMAN. 1990. Syllables and morpheme integrity in Kinande reduplication. *Phonology* 7, 73-119.
- PARK, JAE-ICK. 1997. Disyllabic Requirement in Swahili Morphology. *U. Penn Working Papers in Linguistics* 4.2, 245-259.
- PRINCE, ALAN. 1980. A Metrical Theory for Estonian Quantity. *Linguistic Inquiry* 11, 511-62.
- RIGGLE, JASON and PAMELA MUNRO. 2004. Productivity and Lexicalization in Pima Compounds. *Proceedings of the Berkley Linguistics Society* 30.
- ZURAW, KIE. 2002. Aggressive reduplication. *Phonology* 19, 395-439.