

Reduplication and affixation in Indonesian

Norie Sanchez

The Graduate Center, CUNY

Alan M. Stevens

Queens College and The Graduate Center, CUNY

0. Introduction¹

Reduplication of part or all of a stem as a morphological process is quite common in many of the world's languages, particularly those of Southeast Asia. Indonesian has nasalization processes which interact with reduplication, resulting in apparent overapplication of nasalization rules in reduplicated environments which do not meet the input specifications. This paper will briefly discuss some of the various theories of reduplication which have been put forth and examine the nasalization processes present. An account of Indonesian reduplication which incorporates a three-dimensional framework and prosodic structures will be developed, allowing a clearcut analysis of the reduplication found in this language.

I. Background

Marantz (1982) suggested a model in which reduplication is analyzed as the affixation of a CV skeleton to a stem. The phonemic melody of the stem is then copied onto the affixed CV skeleton and linked to its C and V slots by association rules. Marantz's main points were that reduplication is an affix, the reduplicative affix is represented as a CV-skeleton, the affix is assigned its melody through a process of copying and association, and affixation and copying constitute a single step.

A simple example is found in the Philippine language Agta. The stem **takki** 'leg' is reduplicated to **taktakki** 'legs' in (1).

(1) takki takki takki
 | | | | -> | | | | | | | |
 cvc+cvccv cvc + cvccv = taktakki

Note that the affixed morpheme appears to the left of and on the same tier as the stem, and the unattached melodic units go unrealized.

Marantz's analysis was widely accepted because it handled reduplication like any other affix, and it also automatically ruled out types of reduplication which are imaginable but which do not actually occur in human languages.

Marantz's treatment of reduplication could not account for the apparent over-application of some phonological rules, where a rule appears to apply to both copies of the reduplicated material, although the proper environment is met in only one of the two copies. A case often mentioned in the literature (e.g. Carrier-Duncan (1984)) is found in Tagalog, which contains a rule of Nasal Substitution. A prefix-final nasal combines with the onset obstruent of the stem, resulting in a single nasal segment homorganic with the onset of the stem. For example, when *sayaw* 'dance', is combined with the prefix *maN* (N will be used throughout to indicate the presence of a nasal segment with various surface representations) and what Carrier-Duncan calls R1 reduplication, *maN+R1+sayaw* becomes *ma-na-nayaw* 'dancer.' The Nasal Substitution rule has applied twice -- once where the nasal directly precedes /s/, the correct environment, and once where no nasal precedes the /s/, but *n* appears on the surface.

Using Marantz's model, the derivation would look like (2):

(2) *maN-sayaw* -> *maNsayawsayaw* ->
 | | | | \ | | | | | | | | |
 cv+ cvcvc cv+ cvcvc

ma *na* *sayaw* -> **ma-na-sayaw*
 | | | | | |
 cv+cvcvc

Marantz's solution proposed that nasal substitution in Tagalog is not phonological but morpholexical. The forms **sayaw** and **nayaw** must both be listed in the lexicon; **sayaw** is chosen in certain environments and **nayaw** in other environments, when a "nasal substitution trigger" is present.

Reduplication was further explored in Carrier-Duncan (1984), Clements (1985), Kiparsky (1987), and Mester (1988). Mester's dissertation accounts for reduplication processes in several languages in a three-dimensional framework. Mester's three central hypotheses are that reduplicative templates are morphemes synchronous with the base skeleton, reduplicative templates are directly associated with the base melody (reduplicated forms are thus characterized by a single melody associated with two skeleta), and the linearization of these representations is an instance of Tier Conflation, which takes place at the end of each level. The affix is lined up with the root material according to language- or morpheme-specific rules.

In a three-dimensional framework, any prosodic element may reduplicate, with the reduplicating material on a separate plane from the corresponding plane of the input. This predicts that any phonological rules applying during the cycle in which the reduplicating material is introduced will apply to the string before tier conflation, resulting in an apparent over-application of the rule. Our Tagalog example is repeated in (3), using Mester's framework, with the correct output.

(3) Root: **sayaw** Affix: **maN** + R1, where R1 is **cv**

Cycle 1: Introduction of affix and association to melody:

	cvcvc	Root		cvcvc	Root
maN	sayaw	Melody tier->	maN	nayaw	Melody tier
				\	
cvc	cv	Affix skeleta	cv	cv	Affix skeleta

Tier conflation: **ma-na-nayaw**

II. Indonesian Nasalization Processes

Indonesian is an Austronesian language closely related to Tagalog. Like most of the related languages spoken in the Philippines and the western part of Indonesia, it has a process of nasal substitution, which in this case surfaces uniquely with two prefixes, both of which may combine with reduplication, resulting in apparent rule over-application. Standard Indonesian is referred to throughout.

The two prefixes which show nasalization effects between the morpheme-final nasal of the prefix and the initial segment of the root are the verbal prefix **meN-**, and the nominalizing prefix **peN-**, (where N stands for five alternants: all four Indonesian nasals and zero, i.e. /m, n, ~n/ (written **ny**, except before /c, j/), /ŋ/ (written **ng**) and \emptyset). The reduplication facts which combine with these prefixes will be discussed in section III. The nasalization effects for **meN-** are demonstrated in (4) (the facts for **peN-** are the same).

(4)	<u>stem</u>	<u>meN-form</u>	<u>gloss</u>
A.1.	vowel-stem		
	ajar	mengajar	teach
2.	h-stem		
	hapuskan	menghapuskan	erase
B.	sonorant-stem		
	lalui	melalui	pass by
	rumuskan	merumuskan	formulate
	yakinkan	meyakinkan	convince
	wakili	mewakili	represent
	malukan	memalukan	shame
	nodai	menodai	stain
	nyatakan	menyatakan	state
	ngerikan	mengerikan	blood-curdling
C.	obstruent-stems		
1.	voiced		
	beli	membeli	buy
	duga	menduga	guess
	jaga	menjaga	guard
	gali	menggali	dig
2.	voiceless		
	pukul	memukul	hit

	tulis	menulis	write
	kukur	mengukur	grate
3. affricate			
	cukur	mencukur	shave
4. /s/			
	sapu	menyapus	weep

The prefix nasal surfaces as /ŋ/ (**ng**) before vowel-stems and h-stems and as zero before liquids, nasals and oral glides. Before obstruents, the prefix-final nasal undergoes place assimilation to the following obstruent. A stem-initial voiceless stop or fricative (4C2 and 4C4) is deleted.

Assuming that N is underspecified for place features, we can account for the surface manifestations of the nasal in nearly all cases without further difficulty. The defaults [+back, -cont] for consonants will be inserted at the end of the appropriate cycle, resulting in /ŋ/ in any environment where there are no C-place features available, such as before vowels or segments with no supralaryngeal features at all.

The rule assimilating a nasal to a following obstruent is almost universal in Indonesian. Whether within a morpheme or across a morpheme boundary there are few exceptions to sequences of homorganic nasal plus obstruent. It is crucial to note that the further simplification of these clusters to a single nasal segment occurs **ONLY** with the two prefixes being examined here.

In contexts other than [**meN**+root] we find that nasals delete before identical nasals, as part of a more general degemination rule which deletes the first member of a non-vocalic sonorant geminate, e.g. **garam-mu** 'your salt' -> **garamu**, and **ter-rendam** -> **terendam** 'get soaked.'

Indonesian must have some fully specified nasals underlyingly, since there are instance of nasals which do not assimilate or delete in root contexts. Also, root-final nasals do not assimilate when they appear next to certain suffixes, such as before the suffix **-kan**, eg. **turunkan**, rather than ***turung-kan**, 'to lower' or **yakinkan**, rather than ***yakingkan**, 'convince', or before the phrase-level

enclitic **-kah** 'question particle', e.g. **turun-kah** 'go down?', rather than ***turut-kah** (for these enclitics, see Uhrbach, 1986:11).

Nasal spread is a property only of the two prefixes **meN-** and **peN-**, and this is a highly productive process for these two prefixes. Elsewhere, voiceless stops and affricates do not delete after or assimilate to a preceding nasal, either within a morpheme or across a morpheme boundary. (A few sonorant initial roots do exist where the nasal is not deleted when the prefix **peN-** is added, e.g. **penglihatan** 'sight' from **lihat** 'see'.) Urbach (1986) presents arguments for at least three lexical cycles in Indonesian. Many affixes are not cycle-specific, but may be added on various cycles, which results in an appearance of looping back to an earlier cycle. She assumes that N Assimilation, Obstruent Assimilation and Sonorant Deletion (our Degemination) apply on all lexical cycles. Nasal assimilation is general within the word, but not across word boundaries, thus applying at every lexical level but not post-lexically.

III. Indonesian Reduplication

There are four basic types of reduplication in Indonesian, illustrated in (5). Note particularly the patterns of primary and secondary stress.

- (5)
- a. Reduplication combined with **meN-** prefixation:
memùkul-múkul 'keep on hitting'
 - b. Reduplication of entire word:
búku-búku 'books'
 - c. Reduplication of first consonant:
tetánga 'neighbor'
 - d. Reduplication outside **meN-** affixation:
púkul-memúkul 'to hit each other'

The particular reduplicative configurations which will interest us here combine root reduplication with the nasal prefixes **peN-** and **meN-**: type (5a), where we find consistent overapplication of the nasal assimilation and nasal spread discussed above, and (5d), where we find no overapplication of these rules.

In type (5a) reduplicating forms with the prefix **meN-**, the nasal that results from the rules of nasal assimilation, nasal spread and degemination appears twice -- once in the stem and once in the copy. (6) provides examples in Standard Indonesian.

Notice that overapplication appears on stems that begin with voiceless obstruents; however, in the standard language, there is no overapplication in vowel and h-initial stems.

(6) <u>stem</u>	<u>meN-root-</u>	<u>gloss</u>
ajar	mengajar-ajar	teach
hapus-kan	menghapus-hapuskan	erase
pukul	memukul-mukul	hit
tulis	menulis-nulis	write
kukur	mengukur-ngukur	grate
sapu	menyapu-nyapu	sweep

In (7), a three-dimensional derivation of **memukul-mukul** 'keep on hitting' accounts nicely for the apparent overapplication.

(7) Root: **pukul** Affix: **meN-** + stem reduplication

Cycle 1: Reduplication/affixation:

				c	v	c	v	c	Root
m	e	N	+	p	u	k	u	l	Melody tier
c	v	c		c	v	c	v	c	Affix skeletal

Phonological rules of the cycle: (Nasal assimilation;spread)

					c	v	c	v	c	Root
m	e	m	+		m	u	k	u	l	Melody tier
				\						
c	v				c	v	c	v	c	Affix skeletal

Degemination:

			c	v	c	v	c	Root
m	e	+	m	u	k	u	l	Melody tier
c	v		c	v	c	v	c	Affix skeleta

Tier conflation: **memukul-mukul**

In R2 reduplication, affecting nouns, the entire word is reduplicated, as in (5b). Plurals, for example, are formed by reduplicating the entire singular form. Suffixes such as **-nya** 'his, her, the' are added at a level after the reduplication takes place, since they are never reduplicated. They do, however, trigger stress-shift in the expanded form. Thus, we find **buku-bukunya** 'the books'. The significance of the stress-shift will be discussed below.

R3 reduplication, (example 5c) which is very uncommon, reduplicates the initial consonant of the root, e.g. **tangga** 'stair,' -> **tetangga** 'neighbor.' There are a number of such reduplicated forms, e.g., **sesepeuh** 'elders', from root **sepeuh** 'elder person', and **lelaki** 'male,' from root **laki**, 'husband'. All have a copy of the initial consonant, followed by schwa. (It has been proposed that schwa is epenthetic in other environments in Indonesian (Cohn, 1989:175), and we will assume that it is epenthetic here as well.) This reduplication can be derived very simply, using our three-dimensional framework, as example (8) shows.

(8) Cycle 1: Root: **taNga** Affix: Reduplication of first consonant

c	v	c	c	v	Root
t	a	N	g	a	Melody tier
c					Affix skeleton

Phonological rules of the cycle (Nasal assimilation):

c	v	c	c	v	Root
		/			
t	a	N	g	a	Melody tier
c					Affix skeleton

Tier conflation and schwa epenthesis: **t e t a n g g a**

The final type of reduplication to be considered is that found in (5d), **pukul-memukul** 'to hit each other,' where the reduplication involves a stem prefixed to a morphologically complex word containing a prefix. Trying to derive such forms by either of the models of reduplication outlined so far runs into the following problem: it would be necessary either to order tier conflation before the application of the phonological rules so as to prevent the nasalization rules from applying to the initial /p/ of the reduplication, or we must argue for some means of copying the root and somehow setting it aside where subsequent phonological rules can't interact with it. A third possibility, which has been argued for by Kiparsky (1987:4fn and 98), is that these are root+word compounds. If this reduplication were treated as a prefix, like the other reduplications considered in this paper, it would uniquely be able to precede **meN-**.

In order to account for this type of reduplication, we must rely on prosodic phonology. Cohn (1989) analyzes the stress pattern differences between reduplication and affixation in Indonesian phonological words and clitic groups (as developed in Nespor & Vogel (1986)). We will use Cohn's analysis to determine the order of reduplication/affixation, as well as the presence of one or more than one clitic group.

The phonological word is a prosodic level within the lexicon representing the interaction between the phonological and morphological components of the grammar, and defined as the domain created between left edges of stems [w]. A morphologically complex word may contain more than one phonological word.

A clitic group (C-group) is defined as the domain created by the FIRST left edge, prefix, or stem [c]. According to Cohn,

the expansion to a C-group and application of stress rules first occurs at the end of the first lexical cycle in Indonesian. Addition of any material on a later cycle which triggers an expansion of the C-group forces reparsing and reapplication of certain phonological rules, particularly the Main Stress rule. The addition of a suffix, either lexically or post-lexically, results in formation of a new C-group, triggering reparsing of the word and reassignment of Main Stress. Prefixes, on the other hand, are never stress-bearing or stress-changing in Indonesian, they are never P-words in their own right, and never trigger formation of a new C-Group.

Rather, prefixes are inserted between the leftmost C-group bracket and the first appropriate P-word. (9) shows the parsing of *dicat* 'painted'

(9) root: [c[wcat]] + di- --> [cdi[wcat]]

By using the Indonesian stress facts, particularly the shift of Main Stress triggered by reparsing of an expanded C-group, Cohn shows that reduplication adjoins stems (that is, two P-words, each with its own C-group and main stress) without creating an encompassing Clitic Group. Compounding, on the other hand, results in a single Clitic Group. It is possible to determine whether reduplication or suffixation occurred last, by examining the stress patterns in a reduplicated form. If reduplication has preceded suffixation, the reduplicated material and the suffixed material will reflect different stress patterns, because the expanded C-group will trigger reapplication of the Main Stress rule, resulting in a new across-the-board stress pattern. When suffixation precedes reduplication, not only is the whole C-group reduplicated, but also the stress patterns are the same for both halves of the reduplicated form. Thus, we have *púkul*, 'hit', *memúkul*, 'to hit', and *memúkul-múkul* 'keep on hitting', all showing the expected stress patterns.

The nasal assimilation, nasal spread and degemination processes discussed above are all processes which occur within the P-word, never across the boundary between two P-words or between a P-word and a C-group. If Cohn is correct, they do occur across a single P-word boundary -- that between a prefix and its stem. Cohn's assumption that a

prefix is attached outside the P-word but inside the C-group bracket forces this analysis. If, on the other hand, prefixes were incorporated into the P-word of their host stems, the three phonological rules discussed here would have as their domain the P-word. We will therefore assume that prefixes are NOT outside their host P-words, but rather are incorporated into them, contrary to Cohn. If a reduplicated form contains two C-groups, we would not expect interaction between the adjoining segments at the end of one reduplicated form and the beginning of the next, which is indeed the case.

We now can account for **pukul-memukul**. Although there are compounds consisting of a stem plus an affixed word, the existence of two main stresses in the **pukul-memukul** type reduplications indicate that these are NOT compounds. This reduplication cannot be considered root+word compounding, and we must find some other explanation.

It has already been established that reduplication in Indonesian may occur on various cycles. Also, given the fact that **meN-** is virtually always the outermost prefix, **meN-**forms may be created on later lexical cycles. To account for **pukul-memukul** and other such forms, we need only add a single well-formedness constraint to those proposed in Cohn. If there are two C-groups present in the input to the appropriate cycle, affixes are added to the rightmost Clitic Group. The formation of **pukul-memukul** will be as shown in example (10), where reduplication occurs on the earliest cycle, followed by **meN-** prefixation on the subsequent cycle.

(10) Cycle 1: Root: **pukul** Affix: Reduplication

c	v	c	v	c	Root
p	u	k	u	l	Melody tier
c	v	c	v	c	Affix skeleta

P-word and C-group formation; Stress assignment; Tier conflation: [c[w^púkul]] [c[w^púkul]]

Cycle 2: Affixation to rightmost C-group, application of phonological rules:

[[c v c v c]]					[c[w c v c v c]]					Root		
p	u	k	u	l	m	e	m	u	k	u	l	Melody tier
					c	v						

Tier conflation (No triggering of new C-group, so no change in stress): [c[w_púku_l]] [c[w_memúku_l]]

This well-formedness condition also predicts that we should find more complex forms composed of a left-most Cycle 1 C-group and an expanded, reparsed right-most C-group, which is the case. In the word *hórmát-menghormáti* 'to respect each other', for example, both the prefix *meN-* and suffix *-i* are added to the rightmost C-group, resulting in a stress shift one syllable to the right, but leaving the main stress of the leftmost C-Group unchanged. This morphologically complex word must be formed in the following order: Cycle 1, reduplication; Cycle 2, *meN-* prefixation; Cycle 3, suffixation. The presence of two main stresses was confirmed by our informant.

More complex forms are also possible, including prefixation prior to reduplication, which in turn is followed by *meN-* affixation on the rightmost C-group. Such an example is *sebeláh-menyebeáh* 'side by side' from the root *belah* 'split'. *Sebeláh* 'one half, one side' is formed by prefixing *se-* on the first cycle. On the next cycle, the entire string is reduplicated. Finally, *meN-* is affixed to the rightmost C-group, resulting in the surface form above.

One problem remains to be accounted for. Indonesian also contains words of the form *dipukul-pukulínya*, 'kept on being beaten up by him', with a single main stress, and containing a prefix outside the reduplicated material. Therefore, the entire string must be a single C-group. As just presented, our theory would predict that *di-* should appear on the rightmost *pukul*, assuming two clitic groups.

There is evidence **di-** and **-nya** can be affixed post-lexically. Indonesian allows for the backgrounding of the third person agent in a sentence by incorporation onto the transitive verb. Myhill (1988) and others propose that **di-****-nya** verb forms are actually **di+VERB+agent** with incorporated nominal agents. These forms always have an alternate form **dia + VERB**, as in **dia pukul-pukuli**, '3sg AGENT kept on beating up', where **dia** is a separate pronoun indicating the third singular actor of the sentence. We propose that this **di-** is enclitic to the verb as a backgrounded agent, requiring the **-nya** (AGENT marker), formed in the syntax. The lexically formed passive, in contrast, is **di+VERB** stem with no third person agent marker.

Thus, at the time of noun incorporation in the syntax, the expanded clitic group required for proper formation of **dipukul-pukulinnya** is already present. In the lexicon, on Cycle 2 in (11) a single clitic group is triggered by the suffix **-i**. Post-lexically, the noun incorporation prefix can only attach at the left edge of the unique clitic group present.

(11) Cycle 1 Root: **pukul** Affix: Reduplication

Tier conflation; C-group and stress assignment:

[_c[_wpúkú_l]] [_c[_wpúkú_l]]

Cycle 2 Affixation of **-i**:

[_c[_wpúkú_l]] [_c[_wpúkú_l]]

Input
- i Affix

Tier conflation; Reparsing and Stress reassignment:

[_c[_wpù_kú_l]] [_wpukú_lí_i]]

Cycle 3 Postlexical encliticization of **di-** **-nya**:

[_c[_wpù_kú_l]] [_wpukú_lí_i]]

Input
di- **-nya** Enclitic/Affix

Tier conflation; Reparsing and Stress reassignment:

[_c[_wdipù_kú_l]] [_wpukulí_nya]]

IV. Conclusions

We have provided evidence from Indonesian in support of Mester's conclusions that reduplication is affixation of skeletal material to an existing melody, followed by application of appropriate phonological rules, and ending with Tier Conflation at the end of each cycle. Apparent over-application of phonological rules can be neatly accounted for if we think of reduplication as a three-dimensional process, which allows phonological rules to apply to neighboring segments as is usually the case, before conflation.

We have also explored the complex interactions between stress, affixes and reduplication, in Indonesian by examining the Phonological words and Clitic groups within morphologically complex constructions. Through the formation of such reduplications as **pukul-memukul**, we have been able to show that a morphologically complex word may contain more than one phonological word, but also more than one clitic group. We have also shown that a well-formedness condition on affix placement guarantees that prefixes are placed on the rightmost clitic-group in Indonesian, if two choices are available. Evidence has also been presented for the encliticization of deeply backgrounded third person agents in Indonesian which contrasts with lexical passives with regard to stress and affixation.

Note

¹The authors wish to thank Alex Wijoyo for native speaker intuitions on the data presented here.

References

- Carrier-Duncan, J. 1984. "Some problems with prosodic accounts of reduplication," in Aronoff, M. and R. Oehrle (eds.) *Language Sound Structure*. Cambridge: MIT Press.
- Clements, G. 1985. "The problem of transfer in nonlinear phonology," *Cornell Working Papers in Linguistics*, No. 7.

- Cohn, Abigail C. 1989. "Stress in Indonesian and bracketing paradoxes" *Natural Language and Linguistic Theory* 7:2 167-217.
- Cowper, E. and K. Rice. 1985. "Phonology and reduplication," ms. Canadian Linguistic Association.
- Kiparsky, P. 1987. "The phonology of reduplication," ms. Stanford University.
- Lieber, R. 1980. *On the organization of the lexicon*. MIT doctoral dissertation.
- Marantz, A. 1982. "Re reduplication," *Linguistic Inquiry* 13:435-482.
- Mester, R. Armin. 1988. *Studies in Tier Structure*. New York: Garland Publishing, Inc.
- Myhill, John. 1988. "Nominal agent incorporation in Indonesian," *Journal of Linguistics* 24:111-136.
- Nespor, M. and I. Vogel. 1986. *Prosodic Phonology*. Dordrecht: Foris Publications.
- Prince, Alan S. 1983. "Relating to the Grid", *Linguistic Inquiry* 16: 193-246.
- Sagey, E. 1986. *The Representation of Features and Relations in Non-linear Phonology*. MIT doctoral dissertation.
- Uhrbach, A. 1986. "Levels and Loops in Indonesian." Paper presented at U. of Wisconsin, Milwaukee, Morphology Conference.

