

A MORPHOPHONOLOGICAL ANALYSIS OF -UM- IN JAVANESE

Ekarina

Universitas Katolik Indonesia Atma Jaya
ekarina.winarto@atmajaya.ac.id

Abstract

This study investigates the derivation of words through the process of infixation. Specifically, Prince & McCarthy's (1993a) morphologically-driven Phonological Readjustment theory and Yu's (2007) phonologically-driven Phonological Subcategorization theory are compared to determine which of the two is better fit to explain -um-infixation in Javanese. I show here that Javanese infixation data is crucial in providing insight into this debate, because unlike other previously studied infixation cases, the Javanese -um-always triggers some phonological readjustment(s). This provides overwhelming evidence for the morphologically-driven Phonological Readjustment theory, as this theory is the one that can offer (i) a unified target for infixation in Javanese, and (ii) a phonologically-driven motivation for the derived words.

Keywords: infixation, morphology, phonology, Javanese, syllable structure.

Abstrak

Penelitian ini menyelidiki proses derivasi kata melalui proses infiksasi. Phonological Readjustment dan Phonological Subcategorization merupakan dua teori penting yang berpotensi dapat menjelaskan proses ini. Artikel ini menelaah teori yang lebih cocok untuk menjelaskan data infiksasi -um- dalam Bahasa Jawa, dan data dalam Bahasa Jawa sangat penting dalam perdebatan ini, karena infiksasi -um- dalam Bahasa Jawa selalu diikuti dengan proses penyesuaian fonologis. Hal ini membuktikan bahwa teori Phonological Readjustment dapat menawarkan penjelasan yang lebih natural untuk fenomena ini, karena hanya teori ini yang dapat memberikan (i) target derivasi kata yang terpadu; dan (ii) penjelasan fonologis sebagai motivasi untuk hasil proses derivasi kata dalam Bahasa Jawa.

Kata kunci: infiksasi, morfologi, fonologi, Bahasa Jawa, struktur suku kata.

INTRODUCTION

Affixation is a well-known morphological process that is productive in most languages like English and Indonesian, among many others. Generally, affixation involves attaching a suffix or morpheme to a root and/or a base, to derive (i) new words, or (ii) new grammatical relations. Most commonly, such affixes are attached in front of stems (prefix), or at the end of stems (suffix). Indonesian uses both morphological processes, also in combination, to derive mainly new words. The process summarized in (1) shows a typical prefixation, where the prefix *pe-* is attached before the root *kerja* to form the new word *pekerja*. The process summarized in (2), on the other hand, shows a typical suffixation, where the suffix *-an* is attached after the root *kerja* to form the new word *kerjaan*.

- | | | | | |
|----|------------------|---|------------------|--------------------------|
| 1. | pe- +
prefix | | kerja
to work | → pekerja
worker |
| 2. | kerja
to work | + | -an
suffix | → kerjaan
work |

Often, a new word or new grammatical relations can be built by doing a series of prefixation and suffixation. An example of this is given in (3) below, where the root *beri* is first prefixed with *meN-* to get *memberi* ‘to give. The newly formed *memberi* is in turn suffixed with *-kan* to get the final word *memberikan*, which adds a beneficiary argument.

- | | | | | |
|----|---------------------------------------|---|-----------------|---|
| 3. | a. me-
prefix | + | beri
to give | → memberi
to give (active) |
| | b. memberi
to give (active) | + | -kan
suffix | → memberikan
to give (sth) to (someone) |

It is equally possible, to arrive at the complex verb *memberikan* by first doing suffixation followed by the prefixation – which is probably more plausible if we consider the syntactic bottom-up derivation, but this ordering is not important for the discussion in this paper. We will instead focus on Javanese *-um-*, that is not easily categorized as either a prefix or a suffix.

Most importantly, we have established that prefixation is a process of attaching an affix in front of a root, while suffixation is a process of attaching an affix after a root. In both cases, the affix targets the word boundary, and as such, can be clearly defined. Much less clear is another type of affixation known as infixation. Infixation typically involves inserting an affix in the middle of a word. This phenomenon is not attested in English, with the exception of something familiar that sometimes occur in colloquial speech by “inserting” an expletive or profanity into a word, as seen in (4) below. Usually, the expletive is inserted just before the stressed syllable of the word that it is embedded in.

4. a. abso-**fucking**-lutely
b. un-**fucking**-believable

Another famous set of examples for infixation comes from the Tagalog *-um*. As seen in (5) below, the affix *-um* appears to be inserted after the first consonant in (5a), but the same affix appears to be prefixed in front of the root *inom* in (5b). Similar infixation patterns are widely attested in other languages of the world, and many have attempted to “explain” the rule, derivation and reasons for this pattern we see in (5) (Jiang, 2019; Chan, Lee, & Yip, 2021; Pelagio, et.al., 2022).

- | | | |
|----|-------------------------------------|-----------|
| 5. | a. s- um -ulat
‘to write’ | [Tagalog] |
| | b. um -inom
‘to drink’ | |

Just from comparison of (4) and (5), we can see that infixation rules are language-specific; the pattern in (5) shows furthermore, that within one language, the derivation of the

other hand, gives us a better explanation of why we have infixation, i.e., to maximize the syllable structure in the Tagalog data we have discussed above. *Phonological Subcategorization* does not offer any explanation to why some affixes like prefixes and suffixes sub-categorizes for the morphology (word boundary), while some others like infixes sub-categorizes for phonological entities. However, despite these advantages and disadvantages, both theories are still adequate in explaining the Tagalog data in relation to *-um-*. With the *-um-* data from Javanese, though, I will show that *Phonological Readjustment* is more suitable to explain the data.

METHODS

Poedjosoedarmo's (1979) *Morfologi Bahasa Jawa* was used as the main source for the Javanese infixation data discussed here, because it includes data on the different *-um-* infixation patterns seen in Javanese, including cases where the *-um-* appears as a prefix in vowel-initial base, as well as the variations seen in different base with different initial consonants (see below for more details). Additionally, close friends who are native speakers of Javanese were consulted to confirm the validity of the data taken from Poedjosoedarmo (1979).

JAVANESE DATA

Javanese is an Austronesian language closely related to Tagalog. As such, it is not surprising that the *-um-* affix is also attested in the language. However, the behavior pattern of the *-um-* affix in Javanese is quite different from what we have seen so far from Tagalog. Below, I present this data, that I will use to argue for the *Phonological Readjustment* theory. The Javanese infixation data can be generally divided into two patterns: (i) Pattern A where we see prefixation like our example with the Tagalog *uminom*; and (ii) Pattern B where we see true infixation like our example with the Tagalog *sumulat*.

11. Pattern A:					[Javanese]
(Prefixation + initial vowel /u/ deletion)					
a. Word starting with vowels					
1				2	3 (/u/ deletion)
ili	+	um-	→	um-ili	→ m-ili
'to flow'					
aju	+	um-	→	um-aju	→ m-aju
'to go forward'					
undur	+	um-	→	um-undur	→ m-undur
'to go backward'					
b. Words starting with /l/					
1				2	3(/u/ deletion)
laku	+	um-	→	um-laku	→ m-laku
'to walk'					
layu	+	um-	→	um-layu	→ m-layu
'to run'					

lebu	+	um-	→	um-lebu	→	m-lebu
‘to go in’						
lumah	+	um-	→	um-lumah	→	m-lumah
‘to lie down’						

12. **Pattern B** [Javanese]
(infixation)

a. **Words starting with stops**

1	2	3 (Onset Methathesis)
tibo + um- →	um-tibo	→ t- um -ibo
‘to fall’		
tandaŋ + um- →	um-tandaŋ	→ t- um -andaŋ
‘to do something’		
gantunŋ + um- →	um-gantunŋ	→ g- um -antunŋ
‘to hang’		

b. **Words starting with sibilants**

1	2	3 (Onset Methathesis)
sanak + um- →	um-sanak	→ s- um -anak
‘to be friendly’		
sebar + um- →	um-sebar	→ s- um -ebar
‘to spread’		

DISCUSSION

Compared to the data on Tagalog infixation, the Javanese infixation data provides us with two very interesting facts in Pattern A:

A1. The *um-* “prefixation” pattern applies not only to roots that start with a vowel like *uminom*, but also to roots starting with the consonant /l/ sound.

A2. While Tagalog accepts the less-than-ideal *uminom* (*u.mi.nom*), with the first syllable having no onset, Javanese is stricter in applying the ideal of always having an onset in every syllable. Thus, we get the /u/ deletion pattern observed in (11) above.

The fact observed in (A1) makes it difficult to posit a phonological sub-categorization rule, because unlike in the case for the Tagalog *uminom* and *sumulat*, we cannot define a specific phonological target for the *-um-* in Javanese; the affix seems to target both the first vowel and the first consonant (if it is /l/). Thus, it is more likely that the word boundary is the target for the Javanese *-um-*, giving credit to the theory of *Phonological Readjustment* that prioritizes morphological alignment, before words get readjusted for phonological reason(s). The strict adherence to syllables with onsets seen in (A2), further supports the theory of *Phonological Readjustment*, as we see an obvious phonological motivation for readjustment, specifically to get the ideal phonological structure by way of the /u/ deletion (Step 3); this is an alternative strategy to the *Onset Metathesis* we see happening in Tagalog.

The pattern observed in Pattern B above is similar to the *sumulat* pattern in Tagalog that we have discussed:

B1. When the root begins with consonants other than /l/, the affix -um- surfaces as an infix between the first consonant and the first vowel.

The contrast between roots starting with /l/ and roots starting with other consonants crucially shows that the phonology cannot be easily subcategorized as the target for the -um- affix in Javanese. If we look at words starting with the consonant /l/, we see that the winning output candidate is *mlaku* and not **lumaku* (11b). This means that even though an infixation strategy, similar to the one we see with other roots starting with other consonants is available (Pattern B), it is dispreferred compared to the prefixation. To do this under the *Phonological Subcategorization* theory, we would have to posit a rather complicated and arbitrary phonological sub-categorization, where the target of derivation for the Javanese -um- would have to be either the first vowel or /l/. Even with this arbitrary phonological sub-categorization, we are unable to explain the /u/ deletion observed in Pattern A (11).

In contrast, McCarthy & Prince's (1993a) *Phonological Readjustment* theory offers us a more unified and natural explanation of the data we see in Javanese:

- We can posit the word boundary as a unified target for derivation;
- We can explain both the /u/ deletion in Pattern A, as well as the onset metathesis in Pattern B as phonological readjustment;
- We can also explain the contrast between /l/ and other consonants (discussed further below) as the result of priority ranking for different phonological readjustment strategies.

13. a. Pattern A		[Javanese]														
	<table style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 15%; text-align: center;">1</td> <td style="width: 15%;">um- + ili</td> <td style="width: 10%; text-align: center;">➔</td> <td style="width: 15%; text-align: center;">2</td> <td style="width: 15%;">um-ili</td> <td style="width: 10%; text-align: center;">➔</td> <td style="width: 15%; text-align: center;">3 (/u/ deletion)</td> </tr> <tr> <td></td> <td style="text-align: center;">‘to flow’</td> <td></td> <td></td> <td></td> <td></td> <td style="text-align: center;">m-ili (mi.li)</td> </tr> </table>	1	um- + ili	➔	2	um-ili	➔	3 (/u/ deletion)		‘to flow’					m-ili (mi.li)	
1	um- + ili	➔	2	um-ili	➔	3 (/u/ deletion)										
	‘to flow’					m-ili (mi.li)										
b. Pattern B		[Javanese]														
	<table style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 15%; text-align: center;">1</td> <td style="width: 15%;">um- + tibo</td> <td style="width: 10%; text-align: center;">➔</td> <td style="width: 15%; text-align: center;">2</td> <td style="width: 15%;">um-tibo</td> <td style="width: 10%; text-align: center;">➔</td> <td style="width: 15%; text-align: center;">3 (Onset Methatesis)</td> </tr> <tr> <td></td> <td style="text-align: center;">‘to fall’</td> <td></td> <td></td> <td></td> <td></td> <td style="text-align: center;">t-um-ibo (tu.mi.bo)</td> </tr> </table>	1	um- + tibo	➔	2	um-tibo	➔	3 (Onset Methatesis)		‘to fall’					t-um-ibo (tu.mi.bo)	
1	um- + tibo	➔	2	um-tibo	➔	3 (Onset Methatesis)										
	‘to fall’					t-um-ibo (tu.mi.bo)										

In (13), I show how *Phonological Readjustment* works for the data, whereby the infix -um- targets the left of the root boundary in Step 1 for both patterns, giving us *umili* in (13a) and *umtibo* in (13b) [Step 2]. Then, a phonological readjustment to get only syllables with onsets happen in Step 3, whereby the first vowel is deleted in (13a), while *Onset Metathesis* happens in (13b). The difference in Step 3 strategies is simply due to the availability of onsets in the root.

Now, I will address the difference seen in the phonological readjustment for /l/ and other consonants. If, as illustrated in (13) above, the -um- in Javanese targets the left root word boundary; and the different outcomes between vowel-initial root (13a), and consonant-initial root (13b) is simply a matter of *Onset Metathesis* availability, then why would roots that starts with an /l/ behave differently from other consonants (11 b)? To answer this, I have to make some reference to the *Sonority Sequencing Generalization (SSG)* (Zec, 2007; Bhatta, 2021; Kord, 2018; Wulfert, & Hanulíková, 2022). According to the SSG principle, the vowel nucleus

in a syllable is always the most sonorous – the loudest sound – in the syllable, and the sonority of other co-occurring sounds decrease towards both the left and right tail of the syllable. In other words, if a language allows consonant clusters in the onset position, the first consonant in this cluster would be the least sonorous (least loud), and all subsequent consonants in the cluster would increase in sonority. The opposite is true for codas, with the consonant closest to the vowel being the most sonorous, and the last consonant in the coda cluster being the least sonorous. Figure 1 below shows this sonority hierarchy, i.e., the degree of loudness for different sounds.

VOWELS >> GLIDES >> LIQUIDS >> NASALS >> OBSTRUENTS
 High Sonority ----- Low Sonority

Figure 1. Sonority Hierarchy

In (14) and (15), I have laid out all the possible outputs for the *-um-* derivation involving /l/- initial and other consonant-initial roots. Simple prefixation of *-um-* as in (14a) and (15a) are both not accepted because of the strict adherence to syllables with onsets; this is a pattern we see with the vowel-initial roots too (11a). In the case of *laku*, however, a simple deletion of the initial /u/ is enough to solve this, as seen in (14b). A simple deletion of the initial /u/, as attempted in (15b), though, is not acceptable as the onset consonant cluster **mt* goes against the SSG principle. The /m/ sound is a nasal, and the /t/ sound is an obstruent, so the onset cluster **mt* has a falling sonority, something that is not allowed in the onset position, according to the SSG principle. As such, the grammar is “forced” to apply *Onset Metathesis* as a last resort, so that the output we get still satisfies the requirement for syllables with onset. In the case of *mlaku*, the onset *ml* consists of a nasal and a liquid, so the SSG is satisfied; there is no need to apply *Onset Metathesis*. From this, we can also see that the /u/ deletion as a strategy is preferred, because when both this and the *Onset Metathesis* are available in /l/-initial roots, the /u/ deletion wins over the *Onset Metathesis* in giving us the ideal output in (14b).

14. a. um+laku → *umlaku [um.la.ku] **Syllable without onset** ✗
 b. um+laku → umlaku → mlaku [mla.ku] **SSG satisfied** ✓
 ‘to walk’
15. a. um+tibo → *umtibo [um.ti.bo] **Syllable without onset** ✗
 b. um+tibo → *umtibo → mtibo [mti.bo] **SSG NOT satisfied** ✗
 c. um+tibo → umtibo → tumibo [tu.mi.bo] **Lastresort-Onset Metathesis** ✓
 ‘to fall’

In the following subsection, I provide more evidence for the necessity of phonological readjustment(s) connected to the *-um-* affix.

Bilabial consonant repair

In Javanese, there is further restriction imposed by the grammar of the language with regards to *-um-* attachment. A sequence of bilabial consonants seems to be disallowed in the language, when *-um-* is attached to a base; specifically, there needs to be some repair mechanism if this affix is attached to roots that start with the bilabial consonant /p/ or /b/. The following examples

in (16) show the repair mechanism that is employed by the grammar of the language. Please note that *gemagus* and *keminter* are alternatives to the forms in (16 a & b) in informal speech (Poedjosoedarmo, 1979).

16. **Pattern B** [Javanese]
- a. /b/ → /g/ bilabial → velar
- | | | | |
|------------|------------|-------------------------|-------------|
| 1 | 2 | 3 | 4 |
| um + bagus | → um-bagus | → b-um-agus | → g-um-agus |
| ‘good’ | → | ‘pretending to be good’ | |
- b. /p/ → /k/ bilabial → velar
- | | | | |
|-------------|-------------|--------------------------|--------------|
| 1 | 2 | 3 | 4 |
| um + pinter | → um-pinter | → p-um-inter | → k-um-inter |
| ‘smart’ | → | ‘pretending to be smart’ | |

We see from this data that there is bilabial consonant dissimilation, with the first bilabial consonant in the sequence (/b/ /p/) changing into its velar counterparts, after *-um-*, which has the bilabial consonant /m/, is infix. This is one of the strategies also listed in Zuraw & Yu’s (2009) diverse repairs for multiple bilabial consonants, that may be done to provide contrast between the first consonant and the subsequent bilabial consonant. Consequently, this shows that the process of word derivation involving the *-um-* affix involves several steps and phonological readjustments, after the *-um-* is attached to the root. A simple process of phonological sub-categorization for a specific phonological unit is simply not adequate to explain the Javanese data.

To summarize, the Javanese infixation data that I have presented so far supports McCarthy and Prince’s (1993a) theory of *Phonological Readjustment*, contra Yu’s (2007) *Phonological Subcategorization*, because (i) it is not possible to posit a simple phonological target for the *-um-* affix, and (ii) the patterns we observe necessitates phonological readjustments – to get the ideal syllable structure, as well as satisfy the SSG. The number of syllables in the base does not seem to have any effect on the pattern observed and my analysis.

CONCLUSION

While I have provided an explanation for the different behavior of Javanese *-um-* vis-à-vis consonant-initial roots, /r/ initial roots present a potential problem for the account I have provided here. Since /r/ is categorized as a liquid, just like /l/, SSG predicts that /r/ initial roots should behave like /l/ initial roots (11 – Pattern A). However, the data is messy with some roots surfacing as *r-um-aket* (12 - Pattern B), while some others appear as *m-ringis* or *m-ripat* (11 – Pattern A). The derivations for these words are given in (17) below.

17. [Javanese]
- | 1 | 2 | 3 | |
|----------------|------------------|-------------|-------------------------------------|
| a. um + raket | → um-raket | → r-um-aket | Onset Metathesis – Pattern B |
| | ‘to be familiar’ | | |
| b. um + ringis | → um-ringis | → m-ringis | /u/ deletion – Pattern A |
| | ‘to grin’ | | |
| c. um + ripat | → um-ripat | → m-ripat | /u/ deletion – Pattern A |
| | ‘to fold’ | | |

At this point in time, I am unable to collect enough information on these examples. But there is also some evidence that the Javanese *mlaku* was *lumaku* in Old Javanese. The name of the popular museum located in Yogyakarta, Ullen Sentalu, for example, is short for the phrase *ulating blencong sejatining tataraning lumaku* or “light is the guide for the walk of life,” with the *lumaku* form. This shows that there might be some variation involved in the *-um-* derivation of liquid-initial roots. Nevertheless, I have provided some evidence here that the Javanese *-um-* favors McCarthy & Prince’s (1993a) *Phonological Readjustment* theory. Consequently, the affix *-um-* in Javanese cannot really be categorized as a “true” infix; instead, it is a subset of prefix that targets the left word boundary (Kalin, 2022). What sets it apart from other prefixes is the fact that its attachment to the left of the word boundary triggers other phonological readjustment(s).

NOTE

The author would like to thank two anonymous reviewers who have given constructive feedbacks to the earlier version of this paper.

REFERENCES

- Bhatta, D. D. (2021). Adjacent Consonants and the Universality of Sonority Sequencing Principle in Dotyali Dialects: Syllable Contact Analysis. *Jadila: Journal of Development and Innovation in Language and Literature Education*, 1(3), 254-270.
- Chan, S. S. L., Lee, T. T. M., & Yip, K. F. (2021). Verbal “Infixation” as Partial Deletion: A Case in Cantonese verbs.
- de Almeida, L., Ferré, S., Barthez, M. A., & dos Santos, C. (2019). What Do Monolingual and Bilingual Children with and without SLI Produce When Phonology is Too Complex?. *First Language*, 39(2), 158-176.
- Hale, K. L., & Kenstowicz, M. J. (2001). *Ken Hale: A Life in Language*. Cambridge, MA: MIT.
- Jiang, S. (2019). An OT analysis of in-infixation in Takibakha Bunun. *UST Working Papers in Linguistics*, 11.
- Kalin, L. (2022). Infixes Really Are (Underlyingly) Prefixes/Suffixes: Evidence from Allomorphy on the Fine Timing of Infixation. *Language*, 98(4), 641-682.
- Kord, Z. K. A., Yaghoubi, R., & Malmir, A. (2018). A Study of Sonority Sequencing Principle (SSP) and Syllable Contact Law (SCL) in Three Syllable Words with Initial CVCC in Persian.
- Lee, J. K., & Seo, Y. (2019). A phonetic examination of phonological ambisyllabicity: Focusing on temporal and spectral characteristics. *Linguistic Research*, 36(1), 91-110.

- McCarthy, J., Prince, A., (1993a). Generalized alignment. In: Booij, G., van Marle, J. (Eds.), *Yearbook of Morphology*. Kluwer, Dordrecht, 79–153.
- Pelagio, E. C., Quizon, R., Cruz, J., Atok, J., & Abesamis, K. R. (2022). Kampeangan Perfective Infix < i. In *Papers from the 30th Meeting of the Southeast Asian Linguistics Society (2021)*, 187.
- Poedjosoedarmo, S. (1979). *Morfologi Bahasa Jawa*. Jakarta: Pusat Pembinaan Dan Pengembangan Bahasa, Departemen Pendidikan dan Kebudayaan.
- Rubach, J. (2022). Onset Conspiracy in Upper Sorbian. *Journal of Linguistics*, 1-45.
- Suharno, I. (1982). *A Descriptive Study of Javanese*. Canberra, A.C.T., Australia: Dept. of Linguistics, Research School of Pacific Studies, Australian National University.
- Wulfert, S., Auer, P., & Hanulíková, A. (2022). Speech Errors in the Production of Initial Consonant Clusters: The Roles of Frequency and Sonority. *Journal of Speech, Language, and Hearing Research*, 65(10), 3709-3729.
- Yu, A. (2007). *A Natural History of Infixation*. New York: Oxford University Press.
- Zec, D. (2007). The syllable. *The Cambridge Handbook of Phonology*. Paul De Lacy. Cambridge University Press, 161-194.
- Zuraw, K. and Lu Y. (2009). Diverse repairs for multiple labial consonants. *Natural Language and Linguistic Theory* 27, 197-224.