

November 2022

## Regular sound change; The evidence of a single example

Alexander Adelaar

University of Melbourne, s.adelaar@unimelb.edu.au

Follow this and additional works at: <https://scholarhub.ui.ac.id/wacana>

---

### Recommended Citation

Adelaar, Alexander (2022) "Regular sound change; The evidence of a single example," *Wacana, Journal of the Humanities of Indonesia*: Vol. 19: No. 2, Article 6.

DOI: 10.17510/wacana.v19i2.706

Available at: <https://scholarhub.ui.ac.id/wacana/vol19/iss2/6>

This Article is brought to you for free and open access by the Faculty of Humanities at UI Scholars Hub. It has been accepted for inclusion in Wacana, Journal of the Humanities of Indonesia by an authorized editor of UI Scholars Hub.

# Regular sound change

## The evidence of a single example

ALEXANDER ADELAAR

### ABSTRACT

The Neogrammarians of the Leipzig School introduced the principle that sound changes are regular and that this regularity is without exceptions. At least as a working hypothesis, this principle has remained the basis of the comparative method up to this day. In the first part of this paper, I give a short account of how historical linguists have defended this principle and have dealt with apparent counter evidence. In the second part, I explore if a sound change can be regular if it is attested in one instance only. I conclude that it is, provided that the concomitant phonetic (and phonotactic) evidence supporting it is also based on regularity. If the single instance of a sound change is the result of developments which are all regular in themselves, it is still in line with the regularity principle.

### KEYWORDS

Historical linguistics; sound change; western Indonesian languages.

### 1. INTRODUCTION<sup>1</sup>

Many historical linguists (present author included) believe that the comparative method is still the most reliable tool in comparative-historical linguistics. It is based on the principle that sound changes are regular. In very general terms, this means that if a sound changes, the change happens in every word in

<sup>1</sup> I am grateful to Novi Djenar (University of Sydney), Karl Anderbeck (Universiti Kebangsaan Malaysia), and Eugen Hill (Universität zu Köln), for the very useful comments they gave on a previous version of this paper. They are in no way responsible for any shortcomings in the current version.

---

ALEXANDER (SANDER) ADELAAR is a Principal Fellow in the Asia Institute at the University of Melbourne. His research includes comparative and descriptive linguistics with emphasis on varieties of Malay and the languages of Borneo, Madagascar, and Taiwan. He is the author of *Proto-Malayic* (Canberra: Pacific Linguistics, 1992), *Salako or Badameà. Sketch grammar, texts and lexicon of a Kanayatn dialect of West Borneo* (Wiesbaden: Harrassowitz, 2005), and *Siraya. Retrieving the phonology, grammar and lexicon of a dormant Formosan language* (Berlin: De Gruyter Mouton, 2011). He was co-editor of *The Austronesian languages of Asia and Madagascar* (London: Routledge, 2005) and is currently co-editing the *Oxford Guide to the Malayo-Polynesian languages*. Alexander Adelaar may be contacted at: s.adelaar@unimelb.edu.au.

which the sound occurs. According to the Neogrammarians from the Leipzig School, it happens “without exception”, a claim that has kept linguists arguing for more than a century. Obviously, it needs some explanation, as there are plenty of phonetic developments which at first sight seem to be irregular.

In this paper I will try to provide this explanation with the help of a few short examples. I will also highlight a particular case in which the application of the regularity principle comes into question and which is not usually discussed, namely if a sound change can be shown to be regular even if there is only one example of its occurrence.

Where possible, I take examples from Indonesian languages. I could have used examples from any well-documented language, but Indonesian languages are more likely to be of a direct interest to the readers of this journal. I am also more familiar with Indonesian examples, and I want to promote their use - and the use of Austronesian examples in general - in theoretical discussions of a general historical linguistic nature. These discussions have overwhelmingly been based on data drawn from Indo-European languages, although there are no clear linguistic grounds for this bias towards classical European languages and Indo-Aryan languages. Finally, there has also been a lack of interest in comparative historical linguistics in Indonesia (especially in the comparative method). Hopefully my use of Indonesian examples in this chapter will contribute towards turning this trend around.

The paper is organized as follows. Section 2 clarifies the principle of regularity of sound change and the notion of “exceptionlessness”. Section 3 discusses some new developments in historical linguistics and sociolinguistics that have had some impact on the evaluation of the regularity principle. Section 4 discusses the application of the principle in situations in which the absence of many examples (rather than the existence of apparent counter examples) becomes an issue. Some concluding remarks follow in Section 5.

I obtained the lexical material that I use from the following dictionaries: Stevens and Schmidgall-Tellings (2004) for Indonesian, Wilkinson (1959) for Malay, Pigeaud (1938) for Javanese, and Zoetmulder (1980) for Old Javanese. Ma’anyan data are my personal fieldnotes. Proto Malayo-Polynesian (henceforth PMP) etyma and Proto Western Malayo-Polynesian (henceforth PWMP) etyma are from Blust and Trussel (online) or as indicated. I refer to PMP rather than to Proto Austronesian because the former is closer than the latter to the sample languages I use in this paper, and PMP etyma are easier to interpret than their Proto Austronesian counterparts.

## 2. UNDERSTANDING THE HYPOTHESIS OF REGULAR SOUND CHANGE

Hans Henrich Hock (1991: 35) defines the Neogrammarian regularity hypothesis as follows: “Change in pronunciation which is not conditioned by non-phonetic factors is regular and operates without exceptions at a particular time and at a particular speech community, with possible environmental restrictions. Certain changes (including dissimilation and metathesis) are exempt from this hypothesis”. He explains several of the factors involved in

the definition in order to get a full grasp of the meaning of this hypothesis (Hock 1991: 34-51).

In the first place, regularity and the claim to be without exception only apply to sound changes that take place mechanically and involve pronunciation conditioned by purely phonetic factors. They do not apply to changes that are caused by “the ‘mental’ or ‘psychological’ motivation of other linguistic changes”, such as analogy or borrowing (see below), including (I presume) socially instigated sound change.<sup>2</sup> Sound change is “change of pronunciation which is not conditioned by non-phonetic factors” (Hock 1991: 34).

A recent Indonesian example of the latter would be the tendency to pronounce *a* as *ə* in the suffixes *-kan* and *-an* and the interjection *kan* and in the last syllable of various other roots by some Indonesians during Indonesia’s New Order period (1967-1998). They were emulating some of the speech habits of Indonesia’s head of state at the time, president Suharto. The latter’s Indonesian was heavily influenced by his native language, Javanese, and by the local and informal (Java-Malay) version of Malay used in large parts of Java.

Furthermore, regular sound change usually takes place only (a) at a certain time and (b) in a particular speech community, and it has no currency outside that context.

- a. In the following example, taken from Javanese, the evolution from PMP \**b* to *w* is typically *linked to a certain period in the past*. At some point in time \**b* became Javanese *w*, and it did so in initial word position and between vowels. For instance, PMP \**batu* ‘stone’ became *watu*, PMP \**bulu* ‘body hair; feather’ became *wulu*, and PMP \**qabu* ‘ash’ became *awu*. The change happened a very long time ago, and it also stopped being productive (or “working”) a long time ago. To get a rough idea how long ago that was we may have a glance through Zoetmulder’s Old Javanese dictionary. It both has words in which \**b* has become *w* (showing that the \**b* > *w* change had already had an effect on this historiolect) and words which have *b* (including in word-initial position and between vowels). The words with *b* are not inherited from PMP, so we can safely assume that they came into Old Javanese after the \**b* > *w* had stopped being productive. Apparently, the \**b* > *w* change had already run its course long before Old Javanese became a literary language. And of course, loanwords adopted into *modern* Javanese since the introduction of Islam and the arrival of Europeans have maintained original *b*, compare Arabic *abad* ‘century’ > *abad* ‘idem’, Portuguese *bandeira* ‘flag, banner’ > *bänderə* ‘idem’, and Dutch *kubiek* ‘cubical’ > *kibik* ‘idem’.

<sup>2</sup> Hock is not very explicit about what he categorizes as mentally or “psychologically” motivated change. The inclusion of socially instigated sound change is mine.

- b. That sound changes take place in one particular language and not (necessarily) in others is obvious from the comparison of any two languages that are related. To stick to the current example, the change from PMP \*b to *w* observed in Javanese *watu*, *wulu*, and *awu* did not take place in Indonesian, in which \*b was kept as it was, and we find the corresponding forms *batu*, *bulu*, and *abu*.

Apparent counter examples to the claim of no exceptions are often due to the fact that some changes are conditioned. For instance, if a sound change only happens in a certain position, (for example, at the beginning of a word, or after a stressed syllable) it is still considered to be without exception provided that the condition in question applies. In such a case the condition is part and parcel of the sound change and should of course be accounted for in its formulation, which sometimes it is not. In historical linguistics, a famous case in point is Grimm's Law, which claims that Proto Germanic underwent a consonant shift affecting all its stops. The number of etymological pairs in which the shift is manifested was impressive, but so was the number of exceptions. As it turned out, these exceptions usually happened in clearly marked positions (for instance, if the stop in question was directly preceded by a fricative, or it occurred after a stressed syllable). By reformulating Grimm's Law such that it would take account of these predictable exceptions, and by complementing it with another "law" (Verner's Law specifying that stops occurring after a stressed syllable and in a final syllable undergo different changes), the principle of regularity and absence of exceptions still makes sense.

In the examples from Javanese given above, the change from PMP \*b to *w* is also conditioned, because it only happened *at the beginning of a word or between vowels*. It did not happen in consonant clusters in which \*b was preceded by an \*m, as can be seen in the following examples: PMP \*tambay 'antidote' became *tambo* 'cure'; PWMP \*l̥a(m)baq 'valley' became *lambah* 'idem', PMP \*lumbuŋ 'rice barn, granary' remained *lumbuŋ* 'storage shed, especially for rice, tobacco'. It also did not happen at the end of a word, for example, PMP \*uŋkab 'to open' remained *uŋkab* 'idem', and PMP \*t̥aRəb 'large number' became *tub* 'to be full'.

Obviously, since language is constantly in evolution, it is always possible that a sound change which was totally regular in the past becomes obscured and partly wiped out by a subsequent one, or by lexical borrowing reinstalling the original sound. Taken on face value, the result of such an interplay could easily be interpreted as counter evidence to the regularity of a sound change, as with the loanwords *abad*, *bänderə*, and *kibik* (above) which escaped the historical change from PMP \*b to *w*.

There are also various sporadic changes, such as (a) metathesis, (2) dissimilation, or (3) haplology. Such changes are irregular, even if some phonetics is involved.

- (a) Metathesis occurs when two sounds in a word change position.

An example of metathesis is Indonesian *hidup* 'to live': it developed from an original PMP \**qudip* but the vowels took each other's position in the process. This development will be discussed further in Section 4.

- (b) Dissimilation happens when a sound changes its value because of the presence of a nearby identical vowel.

An example of dissimilation is the Indonesian intransitive prefix *bər-*: it changes to *bəl-* before a root also containing *r*, as in *bər-* + *ajar* 'learning' yielding *bəlarajar* 'to learn'.

- (c) Haplology is seen when in a polysyllabic word two consecutive ones are identical, and one of these identical syllables is deleted.

Classic examples are the term haplology itself being shortened to "haplogy", and morphophonology to "morphology".

Such changes are sporadic: there may be more examples (for example, \**bər-* + *hibur* 'relaxing' *bəlibur* 'to relax, take a day off', Dutch *rapport* > *lapor*), but they by no means take place as a rule: for instance, they are not attested in *bər-atur* 'to line up' and *bər-urus* 'to deal with'. Similarly, no haplology happened in Indonesian *halilintar* 'lightning' (there is no \**halintar*).

Changes due to factors such as (a) analogy and (b) borrowing, which are not phonetically motivated, are also excluded from the definition.

- a. A suitable example of analogy at work are Javanese basic numerals. Although they reflect regular phonetic changes since their evolution from PMP, they have also undergone some changes in order to become disyllabic. Part of them have been disyllabic throughout their history. Other ones reached that state through reduplication, as in the case of *loro*, *papat*, *nənəm*, through an "epenthetic" (added) vowel, as with *ənəm*, and through contraction, as with *siji* and *wolu*, and also with *səpuluh*, which is usually pronounced as [spuluh]:

PMP		Javanese
*sa- 'one' + *biji 'seed'	> *sewiji (+ contraction)	> <i>s-iji</i>
*Duha 'two' > *rwa > ro	> *ro+ro (+ dissimilation)	> <i>loro</i>
*təlu 'three'		> <i>təlu</i>
*həpat 'four' > *(e)pat	> *pat (+ reduplication)	> <i>papat</i>
*lima 'five'		> <i>limə</i>
*ənəm 'six'	> *nəm	> <i>nəm</i>
	(+ reduplication)	> <i>nənəm</i>

	(+ epenthetic /ə/)	>	<i>ənəm</i>
*pitu ‘seven’		>	<i>pitu</i>
*wa-walu ‘eight’	> *wwalu	>	<i>wolu</i>
*siwa ‘nine’	(+ lexical replacement)	>	<i>səŋə</i>
*sa- + *puluq ‘ten’	(+ vowel reduction)	>	[spuluh]

The tendency to become disyllabic is quite general in Javanese roots, but it is even stronger in Javanese counting, in which rhythm plays an important role. Similar tendencies towards length reduction, length uniformity and rhythm are fairly common to languages in general.<sup>3</sup> Note that *ro-* and *pat-* still occur without reduplication in derived numerals, for example, *ro-las* ‘12’, *ro-ŋ-puluh* ‘20’, *ro-likur* ‘22’, *pat-bəlas* ‘14’, *pat-aŋ-puluh* ‘40’, *pat-likur* ‘24’. The only numeral having resisted the tendency towards disyllabicity in free position is *nəm*, but then again, in counting Javanese speakers tend to use *ənəm* or *nənəm* instead. Note incidentally that PMP \*siwa was lost and was replaced by a new word *səŋə*. The origin of the latter remains unknown: however, that has no relevance for our current discussion as it has the same disyllabic word structure as \*siwa.

- b. An example of borrowing interfering with the regularity of sound change can be seen in the following series of PMP etyma and their counterparts (or ‘reflexes’) in Ma’anyan, a Southeast Barito language in Central Kalimantan Province, Indonesian Borneo:

Proto Southeast Barito		Ma’anyan	
*limɛ	‘five’	<i>dime</i>	‘idem’
*kali	‘to dig’	<i>kadi</i>	‘idem’
*kulit	‘skin; hide; rind; bark’	<i>kudit</i>	‘idem’
*lipəs	‘cockroach’	<i>lipas</i>	‘idem’
*litɛ	‘sap of tree or plant’	<i>dite</i>	‘1. sap of tree or plant; 2. sticky rice’
*pili	‘to choose’	<i>pidi</i>	‘idem’
*tali	‘rope’	<i>tadi</i>	‘idem’
*uli	‘to return home; restore; [...]’	<i>udi</i>	‘1. already; 2. go back’

These examples show that in Ma’anyan, Proto Southeast Barito \*l became *d* before \*i, except in *lipas*. The maintenance of \*l in this word may seem like counter evidence to the regularity principle, but there are also strong indications that *lipas* is a loanword. In fact, it shows more irregular sound correspondences: in Ma’anyan, PMP \*ə as a rule became

<sup>3</sup> Schmid (1964: 232) points to the tendency to reduce the length of long numerals in the act of counting.

$\varepsilon$ , and PMP \*s became *h*, although neither of these changes are reflected in *lipas*. The obvious explanation of these irregularities is that *lipas* is borrowed from Malay, which has a corresponding word (*lipas*) with the same form and meaning, and in which the change from \*e to *a* and the maintenance of \*s as *s* are regular. This borrowing explanation is backed up by the occurrence of another Ma'anyan word *dipeh* 'small animal living in freshwater areas', which has a slightly different meaning but is much more in line with the phonological history of the language. While *lipas* should be discarded as a non-inherited vocabulary item, *dipeh* is clearly more "historical" and as such it re-inforces the claim that the sound change from \*l to *d* before \*i is regular.<sup>4</sup>

### 3. POST-NEOGRAMMARIAN DEVELOPMENTS RELEVANT TO THE REGULARITY PRINCIPLE

Some later theories should be mentioned, as they qualify the principle of regular sound change.

Since the 1960s Generativist linguists have been arguing that sound change can also be conditioned by grammar. Nathan Hill (2014) draws attention to their arguments and the case studies they have made in various languages. The case studies include Old Greek, in which Proto Indo-European \*s as a rule is reflected as *h*, except, famously, in those intervocalic positions where it became a marker of aorist or future tense. Hill is able to demonstrate that in each of these cases the apparent evidence for grammatical conditioning can be explained "as some combination of regular sound change, analogy, or borrowing." He also points out that neither the Neogrammarian hypothesis that all sound change is phonetically conditioned, nor the Generativist hypothesis that it can also be grammatically conditioned, can be falsified. Still, to him the belief in regular sound change is more appealing than the idea of grammatical conditioning because "its assumptions are more parsimonious and its descriptive power more subtle".

An example of an apparently grammatically conditioned sound change in Indonesian and Malay is the intransitive prefix *bər-*, which has cognates with initial *m* in other Malayo-Polynesian languages such as Tagalog *mag-*, Toba Batak *mar-*, Malagasy *mi-* and derives from a putative PMP \**may-*. PMP initial \**m* as a rule became Indonesian *m*, and there is no apparent reason why *bər-* should have initial *b*. However, as I will demonstrate below, the *b* in this prefix can still be shown to be the result of a conditioned sound change, given its proximity to \**r* in unstressed antepenultimate syllables.

<sup>4</sup> To be fair, *dipeh* is most probably also borrowed from Malay, but then at a much earlier stage. The evidence runs as follows. PMP originally had \**ipəs* 'cockroach': the addition of *l* is a post-PMP development, which is typical for Malay; since Malay has long been a major influence on Ma'anyan and other West Indonesian languages, it is likely to be the source for irregular forms such as Ma'anyan *lipas* and *dipeh*. That *dipeh* must be a very early loanword from Malay has no bearing on the current discussion, as it has undergone the Southeast Barito change from \**l* to *d* before \**i* in a regular fashion.

A more decisive contribution to the notion of sound change is William Labov's claim that it is also motivated by social factors (Labov 1963 and later publications). Hock (1991: 647) summarizes the conclusions of Labov's research into the pronunciation of the diphthongs *ai* and *au* in the dialect of Martha's Vineyard as follows. (a) Sound change originates in a relatively small number of words and (b) "is generalised to other words in terms of word classes which may be defined phonetically, morphophonemically, morphologically, semantically, syntactically and/or socially (in terms of age group, sex, et cetera)"; (c) "[d]uring the course of this generalization there is a great degree of irregularity and variability", and (d) "regularity is found mainly in the eventual outcome of this change, not in its inception"; (e) [t]he extent of the change correlates to a very large degree with social factors (age, sex, class et cetera); (f) "the extent to which the change is generalised is correlated with social attitude". William Wang (1969) and Cheng and Wang (1975) formulated the concept of lexical diffusion, which is the spread of sound changes through the lexicon of a language. They argue that sound change happens abruptly within a word but spreads only gradually throughout the lexicon. Labov's ideas about socially motivated language change have enriched our insights in the mechanisms of linguistic change considerably, more so than Wang and Cheng and Wang's concept of lexical diffusion. Hock (1991: 649-652) and Hill (2016) argue that neither are in contradiction with the principle of regular sound change. Be it as it may, it has become clear that language change is somewhat more multifaceted than appears from early Neogrammarian discussions. However, as far as its usefulness as a working hypothesis for historical linguistics is concerned, the Neogrammarian notion that sound change is phonetically without exception remains of crucial importance: in fact, without it it is impossible to use the comparative method and make phonological reconstructions.

#### 4. WHAT IF A SOUND CHANGE IS MANIFESTED IN A SINGLE EXAMPLE ONLY?

I would like to present a rather different problem involving the notion of sound change regularity. In straightforward cases, a sound change is manifested in a large number of unrelated words, large enough to claim that the change is regular. But not all cases are straightforward, and sometimes the number of examples is limited. How should we evaluate cases in which there is only one example to illustrate the change in question? (In short, what if there is only one token to demonstrate a certain type of change?). Such a change hardly seems to be "regular". In contrast to other examples seemingly contradicting the principle of regularity of sound change, this instance involves an almost *lack of supporting evidence* rather than the *presence of counterevidence* that needs to be addressed. However, I would like to argue that if it is the result of a configuration of regular sound developments, it is still in line with the principle.

In what follows I discuss three cases of changes for which there is one example

only, as far as I could find. They are taken from Indonesian. One case involves phonetic adaptation to lexical borrowing, whereas two other cases demonstrate a word-structural change and a morphological development respectively. However, all three can be explained as the result of regular tendencies in the phonological history of Indonesian.

#### A. A PHONOLOGICAL INSTANCE

In Malay/Indonesian phonological history, heterorganic consonant clusters were strongly disfavoured, especially in initial position. In modern Indonesian the constraint is no longer in vigour, although much depends on the level of formality and literacy at which the language is used in a given situation. The sequence *d + w* is still rare in Indonesian, except for a large series of words involving the prefix *dwi-* which are mainly the result of a language engineering effort making use of Sanskrit loanwords.

Although Sanskrit, Arabic, and Dutch were important sources for lexical borrowing into Indonesian, none of them abound in *dw* (or *dv*) sequences. Sanskrit has various roots with initial *dv-*, but they often seem to be derived from the root *dvá-* 'two'. Notions like *dvandva* 'pair', *dvīṣ* 'hostility, hatred, dislike', and *dveshya* 'enemy' all involve a binary relation. Arabic does not have initial *dw-*. It has *-dw-* in intervocalic positions, for example, *jadwal* 'schedule',<sup>5</sup> which was borrowed with the same form and meaning into Malay/Indonesian.

Dutch has only a few roots beginning with *dw*, although they occur in various derivations: *dwalen* 'to err, wander', *dwang* 'force, pressure', *dwarrelen* 'whirl, flutter', *dwars* 'cross, thwart', *dwaas* 'silly', *dweil* 'mop', *bedwelmen* 'to drug', *dwerf* 'dwarf', *dwingen* (infinitive), *dwong* (past time) 'to force', and *dwepen* 'to rave, gush about, fanaticise'. One root has intervocalic /*dw*/: *gedwee* 'meek'.

There are only two clear cases of lexical borrowing into Indonesian involving an original initial *dw* sequence:

Sanskrit *dvi-* was borrowed as *dwi-* and used for coining many words involving the meanings 'two, bi-, involving a pair', for example, *dwifungsi* 'dual function'; *dwibahasa* 'bilingual', *dwibulanan* 'bimonthly', *dwisuku* 'disyllabic' et cetera. The introduction of the *dwi-* prefix is part of a language engineering effort by the Indonesian government, aiming at the introduction of new terminology and the replacement of Dutch terminology. While there are many *dwi-* derivations in the dictionary (compare Stevens and Schmidgall-Tellings 2004) and the prefix is generally understood, not all these derivations have found their way into everyday Indonesian. The often heard term with a somewhat infamous connotation *dwifungsi* 'dual function', usually refers to the engagement of members of the military in business deals during Suharto's presidency (1967-1998).

<sup>5</sup> Incidentally, *jadwal* is also odd in Arabic as its four root consonants do not seem to fit the canonical structure of three root consonants in this language.

Dutch *dweil* [dwɛyl] ‘mop’ was borrowed into Indonesian as *pɛl*. In contrast to *dwi-* derivations, this is a clear case of adaptation into the natural spoken language. It is basically the only spontaneous loanword involving initial *dw*. While this seems to be the only instance of a Dutch loanword with an initial *dw*-cluster, I also found another example of a borrowed stop + semivowel cluster but it involves *b + ɣ*. It matches *pɛl* in the sense that the resulting loanword also shows cluster reduction with the surviving semivowel having undergone fortition. Compare Dutch *object* [obʲɛkt] ‘object’: in official Indonesian it is matched by the somewhat formal sounding *obʲɛk* which basically has the same meaning, although it also has the connotation of an object or means to make money on the side, to moonlight, use the company’s motor vehicle as a taxi. In everyday spoken language it became *ojek* ‘bicycle or motorcycle put to use as an inexpensive means of hired transportation’, and *η-ojek* ‘to transport a paying passenger on a bike or motorcycle’.

Semantically, matching Dutch *dweil* with Indonesian *pɛl* causes no problem. Phonetically, it involves a change from *dw* to *p*. I assume that the change is regular, although the supporting evidence is minimal, with only this one attested case.

#### B. A PHONOTACTIC INSTANCE

PMP \**qudip* ‘to live; alive’ became Indonesian *hidup* ‘idem’: the sound changes in it are regular, but the metathesis of vowels is not. And while there are many Malayo-Polynesian languages that have a reflex of \**qudip*, the metathesis is typical for Indonesian and other Malayic languages. In fact, it is a defining feature of Malayic, which is a genetic linguistic subgroup. It is a strong argument for including Iban, Banjarese (in Borneo), Minangkabau, Kerinci (in Sumatra), Kelantan Malay (in West Malaysia) and Urak Lawoi’ (in southern Thailand) in this subgroup, and for excluding Lampung (in Sumatra) and Embaloh and Bidayuh (in Borneo) from it (Adelaar 1992).<sup>6</sup>

What motivated this metathesis? Although there is currently no constraint against final \*-*ip* and \*-*im* sequences in the word structure of Malayic languages, there must have been one historically.

Indonesian has various words that have *i* in the last syllable and end in a labial consonant, such as (*h*)*intip* ‘to spy, lurk’, *kacip* ‘betel nut scissors’, *kədip*, *kəlip* ‘blink, flicker’, *kətɪp* ‘nip or bite (of small insects)’, *kutip* ‘to quote’, *lancip* ‘smooth and pointed’, *nasib* ‘fate’, *sirip* ‘fin’, *sisip* ‘to insert’, *cicip* ‘to taste’, *titip* ‘to entrust’, *kilim* ‘seam’, *hakim* ‘judge’, *Muslim* ‘Muslim’, *kirim* ‘to send (object)’. However, it does not take long to see that many of these words are of Middle Eastern extract: *nasib*, *hakim*, and *Muslim* are originally Arabic, and *kilim* derives from a Central Asian *kelim* ‘kind of rug’ and has its ultimate origin in Persian

<sup>6</sup> There are, however, a few doubtful cases such as Sundanese, which has *hurip* ‘to revive, flourish’ and *hirup* ‘living, alive’, Balinese, Sasak, and Rejang *idup* ‘to live’.

*gilim* ‘garment made of wool or goat hair’. The sources for *cicip*, *kacip*, *kəlip*/*kədip*, *lancip*, *sisip*, *cicip*, *titip* are not immediately clear although these words also occur in Javanese, which has exercised a major influence on Malay for more than a millennium and has no constraint against *i* followed by a final labial. Moreover, *titip* is labelled as originally Javanese in both Wilkinson (1959) and Klinkert (1916), and *sisip* has a variant *sisit* (Wilkinson 1959). The word *kirim* has a more established history and cannot readily be explained as a loanword. It has cognates with the same meaning in Javanese (*kirim*), Sundanese (*kirim*), Madurese (*kèrèm*), and various Malayic languages; however, note Iban, which shows a change from \**i* > *u* in one of its corresponding variant forms *kirim* and *kirum* (both same meaning). Note also that all these cognates of *kirim* have identical meanings: this would be somewhat unexpected if they were related through common inheritance, but makes more sense if they were related through borrowing. Last but foremost, most Indonesian words ending in *-ip* or *-im* lack PMP etyma, which underscores their historical shallowness.<sup>7</sup>

In conclusion, it is not possible to explain all Indonesian lexicon ending in *-ip* and *-im* as loanwords. However, most Indonesian words ending in *-ip* or *-im* have a rather shallow history. And conversely, the few PMP etyma ending in \**-ip* or \**-im* either lack an Indonesian reflex or, in the case of \**qudip*, underwent vowel metathesis. This suggests that in Indonesian/Malay history, the *u/i* metathesis is due to a phonotactic constraint against high front vowels preceding labial consonants in last syllables. If so, the constraint is manifested in only one Indonesian example; at the level of the Malayic subgroup, the doublet form *kirum* alongside *kirim* in closely related Iban is another manifestation.

Finally, by making its motivation explicit, the constraint also explains why the supposed metathesis has occurred. It makes the metathesis more “regular”, showing that it is ultimately triggered by Malay phonetics, in spite of the fact that the notion of metathesis generally does not meet the Neogrammarian regularity requirement and is only seen as a tendency.

### C. A MORPHOLOGICAL INSTANCE

The PMP agent-oriented verb prefix \**may-* became the Indonesian (Malay) intransitive verb marker *bər-* (Adelaar 1992: 163). This is not an obvious development: while the semantic evolution of *bər-* can be accounted for, its shape is puzzling. The change from \**a* to *ə* is regular (see below), but the change from \**m* to *b* seems not. On face value, it supports the Generativist view that sound change can be grammatically conditioned. However, a closer look reveals that there is a perfectly phonetic explanation for the denasalisation of \**m* given the word structural setting in which the change took place.

<sup>7</sup> Blust and Trussel (online) do not include comparative data for Malay *kirim*. Dempwolff (1938) has \**kirim* ‘to send’, but the evidence is very weak, with reflexes in Javanese and Malay only.

In Indonesian, as in Javanese and most other Malayo-Polynesian languages, disyllabicity is the favoured root structure (Blust 2013: 213, 234).

Furthermore, in trisyllabic roots, the antepenultimate syllable has schwa as a default vowel, for example, *bənua* ‘continent’, *tətaŋga* ‘neighbours’, *gəmuruh* ‘rumble’. Regular exceptions to this are roots in which the antepenultimate vowel is followed by a weak consonant: these roots will have *a* before *h*, *u* before *w*, and *i* before *y*, for instance, *dahulu* ‘past; first’, *suatu* ‘one’, *biyawak* ‘lizard’. Some of these words also has a short variant form in which the first syllable is lost, for example, *dulu* ‘in the past’, *satu* ‘one’. The situation described here is the result of “antepenultimate vowel neutralisation”: whatever the antepenultimate vowel was historically, it will regularly end up as a schwa unless it takes on the colouring of a following \*h, \*w or \*y; *dahulu* developed from \*di \*hulu ‘at the head/beginning’, *suatu* developed from \*sa- ‘one’ and \*batu ‘stone’, and *biawak* developed from PWMP \*wayawak (Blust and Trussel online). Note that in the latter cases the resulting antepenultimate vowel most likely also became a schwa at some intermediate stage, for example, \*di \*hulu > \*dəhulu > *dahulu* (also *dulu*); \*sa-batu > \*sawatu > \*səwatu > *suatu* (also *satu*); \*wayawak > \*wəyawak > \*wiyawak > *biawak*.

There are also other – irregular – exceptions, which include loanwords, such as *wanita* ‘woman’, which is originally Sanskrit, or *boneka* ‘puppet’, from Portuguese, and some rare cases which remain historically unexplained, such as *binatan* ‘animal’.<sup>8</sup>

This antepenultimate vowel neutralisation and subsequent vowel colouring before weak consonants is most probably a consequence of the relative lack of stress on antepenultimate syllables. Indonesian admittedly has no distinctive word stress, and PMP was also reconstructed without it. Moreover, Indonesian word stress is hardly perceptible. However, as far as it is perceptible, it is on the penultimate syllable of a root unless the latter has a schwa, in which case it is on the last syllable. Suffixation brings about a shift to the next syllable (Sneddon et al. 2010: 11-12).

In PMP, non-verbal roots could become verbs through prefixation of \*paN- (+ distributive) or \*paR- (+ durative), with \*maN- and \*maR- respectively as active voice counterparts. These prefixes survive in many western Malayo-Polynesian languages, although they often have changed their meaning and function. Concentrating on the active voice prefixes, observe the following examples:

<sup>8</sup> This word consists of a PMP root \*batan ‘trunk, stem; body; self [...]’ and \*<in>, which in some languages (for example, Batak) is a nominal infix. Its meaning is historically suspect, as PMP does not seem to have had a generic term for ‘animal’, and terms for this notion in current Malayo-Polynesian languages cannot be traced to a PMP etymon with the same meaning. Furthermore, the derivation of nouns with \*<in> is not typically Malayic, suggesting that *binatan* was originally a loanword.

	*maN-	*may-
Tagalog (Philippines) <sup>9</sup>	maN- (+ active voice)	mag- (+ active voice)
Iloko (Philippines) <sup>10</sup>	maN- (+ de-transitive)	ag- (+ intransitive)
Malagasy (Madagascar) <sup>11</sup>	maN- (+ active voice)	mi- (+ active voice)
Toba-Batak (Sumatra) <sup>12</sup>	maN- (+ active voice)	mar- (+ active voice)
Karo-Batak (Sumatra) <sup>13</sup>	N- (+ active voice)	ər- (+ intransitive)
Indonesian (Sumatran origin) <sup>14</sup>	məN- (+ active voice)	bər- (+ intransitive)

The vowel change in Indonesian *bər*-is regular: given the language's predisposition to use disyllabic roots, prefixes are almost always in antepenultimate position (\*may- + \*CVCVC). However, the change from initial \*m to *b* does not seem regular because in other cases, Indonesian \*m as a rule remains *m*.

One can think of the following explanations for *bər*- :

- It does not reflect \*maR- because it looks different and marks intransitivity; it also occurs in various other languages in Sumatra and Borneo which have *bə(r)*-, *ba(r)*-, *bara*-, *b*- et cetera marking intransitivity (including in Ngaju Dayak and Ma'anyan Dayak).
- It reflects \*maR-, but the fact that it has *b*- instead of expected *m*- agrees with the observation that affixes do not always undergo the same sound changes as the lexicon. It would affirm the Generativist view that sound change can also be grammatically conditioned.
- It reflects \*maR-, but the phonetic conditioning of the change from \*m to *b* is not straightforward and has to be established first.

As far as the first explanation is concerned, note that the non-Malayic languages in Sumatra and Borneo that have *bə(r)*-, *ba(r)*-, *bara*-, *b*- et cetera have all borrowed heavily from Malay; moreover, Ma'anyan *ba*- occurs side by side with *mi*-, which is the regular reflex of \*maR-. In these languages, the borrowed status of the prefix is generally transparent.

In Adelaar (1992: 163) I argue for the third explanation based on the following considerations:

- The antepenultimate syllable precedes the stressed syllable.
- In this relatively unstressed syllable the vowel becomes reduced and is neutralised to a schwa. (In fast speech this schwa even becomes  $\emptyset$  when

<sup>9</sup> Himmelmann (2011: 363).

<sup>10</sup> Rubino (2011: 336).

<sup>11</sup> Rasoloson and Rubino (2011: 472-473).

<sup>12</sup> Van der Tuuk (1971: 92-114).

<sup>13</sup> Woollams (2011: 542-545).

<sup>14</sup> Adelaar (1992: 149-150, 155-156).

it is followed by a liquid, for example, *bəraŋkat* ‘to leave’ and *bəlar* ‘to study’ (Section 2) are usually pronounced [braŋkat] and [blajar] respectively).

- c. As a result of this reduction, \*m and \*r come close together and tend to form a cluster, which is disfavoured in the phonotactics of Malayic languages.
- d. An epenthetic *b* emerges as a result: \*may- > \*mər- > \*m<sup>ə</sup>r- > \*mb<sup>(ə)</sup>r-;
- e. As initial consonant clusters are not tolerated in the phonotactic structure of Malay, \*mb<sup>(ə)</sup>r- is reduced to *b<sup>(ə)</sup>r-*, which is the current intransitive prefix with its various allomorphs *b<sup>ə</sup>r-*, *b<sup>ə</sup>-*, *b<sup>ə</sup>l-*, and (in fast speech:) *br-* and *bl-*. (Note that phonologically the shortness of the vowel in *bər-* is irrelevant, as length is not distinctive in Malay).

In summary, the change from \*may- to *bər-* is not irregular but it is the result of an unusual concatenation of phonological and prosodic circumstances that are regular in themselves. There is no need to appeal to the Generativist theory that sound changes can also be grammatically conditioned, or to a total rejection of any historical connection. There is only one other Malay prefix in which initial \*m comes in the direct vicinity of a liquid, namely the PMP plural and reciprocal/reflexive prefix \*maR-si-, which became *bərsə-*, an infrequent and unproductive reciprocal marker still observed in some forms such as the following:

<i>tubuh</i> ‘body’	<i>bərsə-tubuh</i> ‘to fornicate’
<i>tumpu</i> ‘take-off, abutment’	<i>bərsə-tumpu</i> ‘take off against each other (in tug-of-war game)’
<i>mənakan</i> ( <i>təkan</i> ) ‘to press’,	
<i>lutut</i> ‘knee’	<i>bər-si-təkan lutut</i> ‘with one’s arms (leaning) on one’s knees’

It is possibly also observed in *bərsəlisih* ‘1. to fight; 2. (obsolete) to pass in the night (of boats on a river)’, although that form is also open to a different analysis.<sup>15</sup> The prefix *bərsə-* seems to have developed under the same combination of phonetic and phonotactic conditions as *bər-*. Most likely, its development was not separate from that of the latter, and it should be analysed as a combination of *bər-* and *sə-*.

In fact, there are more examples of epenthetic homorganic stops between a nasal and a following liquid (both *l* and *r*), but they usually emerge in the penultimate syllable of trisyllabic loanwords instead of prefixes in

<sup>15</sup> That is, it could also be analysed as a prefixed verb *bər-səlisih*, the root of which derives from a historical \*sisi(h) ‘edge’ with the fossilised infix \*<əl> expressing diffusion.

antepenultimate position. Observe how the following loanwords became adapted (Adelaar 1988: 65):

Portuguese	<i>inglês</i> (Colloquial [inrê]) 'English'	>	Indonesian <i>ingəris</i> 'English'
Dutch	<i>kameraad</i> 'comrade'	>	Malay (obsolete) <i>kambrat</i> (? [kambərat]) 'comrade'
English	<i>general</i>	>	Indonesian <i>jenderal</i> [jɛndəral]
Arabic	<i>jumlah</i> 'sum, whole, body, crowd'	>	Jakarta Malay <i>jumbəle</i> 'number, quantity'
(Spoken) Arabic	<i>bəsməllæ:h</i> 'in the Name of God (uttered during ritual slaughter)'	>	+ <i>bə-sm(ə)ləh</i> + back-formation > (originally) Malay <i>səmbələh</i> 'slaughter according to Muslim prescription' > Indonesian <i>səmbəlih</i> 'to kill, slaughter'

Consider also the following disyllabic loanword:

Dutch <i>emmer</i> ([ɛmər])	>	Indonesian <i>ember</i> ([ɛmbər])
-----------------------------	---	-----------------------------------

And the following inherited root in Minangkabau, a Malayic language:

Proto West-Malayo-Polynesian * <i>timəraq</i> 'tin foil'	>	<i>timbarah</i> (via an intermediate stage * <i>timbərah</i> ) 'tin foil'
---	---	--

These examples clearly illustrate the likelihood of epenthetic homorganic stops to emerge in environments consisting of an initial nasal directly followed by a reduced schwa and a liquid. The prefix *bər-* is the only such instance in word-initial position. There is no reason to consider it "irregular" given the fact that there are no other prefixes the initial consonant of which became denasalized.<sup>16</sup>

## 5. CONCLUDING REMARKS

Each of the changes discussed in Section 4 are unique in that there is only one instance to demonstrate the proposed analysis. A change that is manifested in one instance only is difficult to defend and should usually be rejected due to a lack of evidence. However, an investigation of the pathways they followed in their development shows that they are regular. In the case of *pəl*, there is compelling evidence based on word structural and extralinguistic (cultural) evidence. In the case of *hidup*, there is strong word structural evidence in historical hindsight. Finally, in the case of *bər-* the evidence is purely phonetic, taking into account that the change took place in an under-stressed syllable.

<sup>16</sup> Malay *di-* is sometimes explained as reflecting PMP \**ni-* and having a denasalized initial \**n*; see Adelaar (2005, 2009) for an alternative explanation.

This created an environment in which a nasal and a following liquid came into sufficiently close proximity to cause the emergence of an epenthetic \*b. The sequence of changes leading from \*maR- to bər- are all regular in themselves, even if the final outcome is unexpected and there are no analogous cases to make them seem more regular.

## REFERENCES

- Adelaar, Alexander. 1988. "More on Proto-Malayic", in: Mohd. Thani Ahmad and Zaini Mohamed Zain (eds), *Rekonstruksi dan cabang-cabang Melayu Induk*, pp. 59-77. Kuala Lumpur: Dewan Bahasa dan Pustaka.
- Adelaar, Alexander. 1992. *Proto-Malayic; The reconstruction of its phonology and parts of its lexicon and phonology*. Canberra: Australian National University, Research School of Pacific Studies. [Pacific Linguistics Series C-119.]
- Adelaar, Alexander. 2005. "Much ado about di-", *Bijdragen tot de Taal-, Land- en Volkenkunde* 161/1: 127-142.
- Adelaar, Alexander. 2009. "The many origins of the passive prefix di-", in: A. Adelaar and A. Pawley (eds), *Austronesian historical linguistics and culture history; A festschrift for Robert A. Blust*, pp. 129-141. Canberra: Pacific Linguistics.
- Adelaar, Alexander and Nikolaus P. Himmelmann (eds). 2011. *The Austronesian languages of South East Asia and Madagascar*. Second edition. London: Routledge, Taylor & Francis Group. [Routledge Language Family Series; First edition 2005.]
- Blust, Robert A. 2013. *The Austronesian languages*. [Asia-Pacific Linguistics, 008. Australian National University.]
- Blust, Robert A. and Steven Trussel. Online. *Austronesian comparative dictionary*. Honolulu: Department of Linguistics, University of Hawai'i at Mānoa. [Http://www.trussel2.com/acd/.]
- Cheng, Chin-Chuan and William S-Y Wang. 1975. "Sound change; Actuation and Implementation", *Language* 51: 255-281.
- Dempwolff, Otto. 1938. *Vergleichende Lautlehre des austronesischen Wortschatzes III: Austronesisches Wörterverzeichnis*. Berlin: Reimer. [Zeitschrift für Eingeborenen-Sprachen Beiheft 19.]
- Hill, Nathan W. 2014. "Grammatically conditioned sound change", *Language and Linguistics Compass* 8/6: 211-229.
- Hill, Nathan W. 2016. "A refutation of Song's (2014) explanation of the 'stop coda problem' in Old Chinese", *International Journal of Chinese Linguistics* 3: 270-281.
- Himmelmann, Nikolaus P. 2011. "Tagalog", in: A. Adelaar and N.P. Himmelmann (eds), *The Austronesian languages of South East Asia and Madagascar*, pp. 350-376. London: Routledge, Taylor & Francis Group.
- Hock, Hans Henrich. 1991. *Principles of historical linguistics*. Berlin / New York: Mouton de Gruyter.
- Klinkert, H.C. 1916. *Nieuw Maleisch-Nederlandsch woordenboek*. Third revised edition. Leiden: Brill.

- Labov, William. 1963. "The social motivation of a sound change", *Word* 19/3: 273-309.
- Pigeaud, Th.G.Th. 1938. *Javaans-Nederlands handwoordenboek*. Groningen: Wolters Noordhoff. [Second edition 1982; Leiden: Royal Institute of Linguistics and Anthropology.]
- Rasolosoan, Janie and Carlo Rubino. 2011. "Malagasy", in: A. Adelaar and N.P. Himmelmann (eds), *The Austronesian languages of South East Asia and Madagascar*, pp. 456-488. London: Routledge, Taylor & Francis Group.
- Rubino, Carlo. 2011. "Iloko", in: A. Adelaar and N.P. Himmelmann (eds), *The Austronesian languages of South East Asia and Madagascar*, pp. 326-349. London: Routledge, Taylor & Francis Group.
- Schmid, Heinrich. 1964. "Zur Entwicklungsgeschichte der romanischen Zahlwörter", *Vox Romanica* 23: 186-238.
- Sneddon, James N., Alexander Adelaar, Dwi Noverini Djenar, and Michael C. Ewing. 2010. *Indonesian; A comprehensive grammar*. Second edition. London: Routledge.
- Stevens, Alan M. and A.Ed. Schmidgall-Tellings. 2010. *A comprehensive Indonesian-English dictionary*. Athens, OH: Ohio University Press. [First edition 2004.]
- Tuuk, Herman Neubronner van der. 1971. *A grammar of Toba Batak*. The Hague: Martinus Nijhoff.
- Wang, William. 1969. "Competing changes as cause of residue", *Language* 45: 9-25.
- Wilkinson, R.J. 1959. *A Malay-English dictionary (Romanized)*. London: Macmillan.
- Woollams, Geoff. 2011. "Karo Batak", in: A. Adelaar and N. P. Himmelmann (eds), *The Austronesian languages of South East Asia and Madagascar*, pp. 534-561. London: Routledge, Taylor & Francis Group.
- Zoetmulder, P. 1980. *Old Javanese-English dictionary*. Two Parts. The Hague: Nijhoff. [KITLV publication.]