Odd conditions

Context-sensitive sound change in unexpected contexts

Robert Blust
University of Hawai'i

Since the 19th century linguists have expected to find conditioned sound changes in environments that make phonetic sense: consonants palatalize adjacent to front vowels, back vowels front if a front vowel occurs in the next syllable, stops voice between voiced segments, and so forth. Most conditioned sound changes conform to this expectation, but a surprising number do not. Some of these are well known, as the palatalization of *s before most word-initial consonants in High German. Since there is no obvious explanation for them, such changes are generally ignored in discussions of historical phonology. The result of this practice has been to give the false impression that what appear to be phonetically unmotivated sound changes are rare abnormalities that probably would conform to expectation if we had more information about them. This paper draws attention to examples of conditioning in Austronesian languages in which the phonetic properties of the context appear unrelated to those of the change, and it questions why such changes should occur. Although finding a completely satisfactory explanation has proven difficult, one general conclusion suggested by the data is that native speakers have an intuitive recognition of natural classes that is independent of phonetic motivation.

Keywords: sound change, conditioning, phonetic motivation

1. Introduction

In an earlier publication (Blust 2005) I drew attention to a number of surprising innovations in Austronesian languages, and concluded that few if any of them can convincingly be explained as due to previously unknown phonetic mechanisms, or to a telescoping of more typical changes over time. This has not stopped some scholars from continuing to speculate about possible phonetic conditions

Journal of Historical Linguistics 7:3 (2017), **322–371**. DOI 10.1075/jhl.15013.blu
ISSN 2210–2116 / E-ISSN 2210-2124

that apply very rarely, or about transitional steps that have never been observed (e.g. Beguš 2015). As a science, linguistics is naturally driven to explain individual observations as consequences of the operation of general principles, and on the whole, this has been highly successful. However, it does not follow from the inductive logic of science that because a research program has accounted for observations *a, b, c* it can be extended to account for observations *d, e, f* on the basis of a principle-driven account alone. To echo Hume, because all swans in one's experience have been white can never justify declaring "all swans are white". Scientific theories are rooted in observation, and although they may reach levels of abstraction that seem remote from their observational bases, any argument based on *a priori* plausibility should be disfavored when compared with an alternative that demands falsifiable evidence.

Needless to say, my purpose in discouraging arguments based on speculation is not to be anti-explanatory, but rather to insist that explanations be empiricallly grounded. To choose an example from my earlier paper, *w, *b and *mb are recurrently reflected as *c-*, *-nc-* ([ʧ], [nʧ]) in Sundanese, a major Austronesian language of west Java that has been studied by linguists for well over a century. To overcome the featural absurdity of such an apparent change one might speculate that this situation came about through a series of innovations *b > *w > *y > *c-*, *-nc-*. However, since *y remained unchanged, and *-mb- became *-nc-* in at least two known cases, it quickly becomes difficult to find a way to make this work. With enough persistence and freewheeling imagination, some contrived explanation can no doubt be found. But my point in trying to raise awareness about the importance of such "unnatural" cases is that an acceptable explanation (i.e. one based on phonetic motivation) can usually be salvaged only by proposing a series of steps for which no evidence exists. Rather than "whitewash" such apparent changes or hide them in a closet of silence, it seems much more in the interest of science to recognize that they challenge a widely-held assumption, namely the assumption that because *most* sound changes are phonetically motivated *all* sound changes are phonetically motivated. In short, only by incorporating deviations from expectation in a theory of sound change can we hope to gain a deeper understanding of why expected changes work the way they do.[1]

1. Janda & Joseph (2003: 207) point out correctly that the term "sound change" is often used without making the critical distinction between "innovation" and "diachronic correspondence", or what Blust (2005: 226) called "sound change" and "reflex", and they remind the reader of facts that have long been known in Germanic linguistics, namely that changes which are phonetically motivated at the outset may later be generalized to apply in environments where the phonetic conditions needed to produce them never existed, as with High German umlaut. Although it is discursively more convenient to speak of "changes" than of "reflexes" when the distinction is in fact at issue, I have endeavored to keep this difference in mind at all times in considering

One reason for revisiting this topic is because deeply-entrenched beliefs are such that even scholars who regard themselves as open-minded are likely to quickly forget a single article or statement that challenges their belief system, and move on with "science as usual". However, another reason is connected with the difference between the content of a sound change and the environment in which it occurs. In principle, one can *always* speculate that a change *x > y took place through intermediate steps *x > x_1 > x_2 > x_3 > y, but it is far more difficult to propose such chains of speculation in relation to the environment of a conditioned sound change, which is commonly conceived as a fixed matrix of phonological features. With this in mind I have assembled a small number of reflexes in a wide sample of Austronesian languages which are uncontroversially seen as conditioned sound changes. In most cases the innovation itself, unlike the cases reviewed in Blust (2005), is a garden-variety sound change. What raises scholarly eyebrows is the *environment* that appears to condition the change, in particular the absence of any clear featural relationship between change and condition. In some cases, these changes are unique to individual languages, but in others they are recurrent, and it is the latter which present the greatest explanatory challenge. On the one hand, if such changes are phonetically motivated, why do they not occur with a globally scattered distribution in the world's languages like, say, palatalization before front vowels, or final devoicing? On the other hand, if they are not phonetically motivated, why have they occurred repeatedly through what appear to be historically independent events that are nonetheless restricted either by geography or by established genetic relationship?

The further organization of the paper is as follows. First, I consider a textbook example of a problem in phonological reconstruction that is virtually impenetrable until the reader recognizes a type of phonological conditioning that is totally unexpected. Having thus set the stage for what is to come I then examine six case studies of vowel change in which there is no theoretically recognized relationship between the phonetic properties of the segment undergoing change and the properties of the conditioning environment. After this I consider six consonant changes that show the same lack of correspondence between the featural composition of the affected segment, and that of the phonetic matrix in which it is embedded. Finally, I present my conclusions.

---

the following data. However, where the telescoping of proposed innovations is a matter of pure speculation, I avoid it as contributing nothing to our understanding of the data (with regard to this point cf. further Garrett 2015: 232–233).

2. Aphonetic conditioning of sound change?

Most sound changes have transparent phonetic motivations, and this is particularly evident with conditioned change. However, some conditioned changes appear to operate regularly in environments that have no obvious causal connection with the phonetic adjustments of the segment undergoing change. This observation is not new, and is sometimes immortalized in linguistic terminology, as with the mysterious *ruki*-rule of Sanskrit, but phonetically opaque sound changes are commonly shunted aside because they violate the general expectation that conditioned changes tend to be products of assimilation. The Second Germanic Consonant Shift, formulated by Jacob Grimm in 1822, distinguished a number of phonological differences between the more conservative Low German dialects and the more innovative High German dialects, including several types of palatalization. Most of these have transparent phonetic motivations, but this is not true for the palatalization of *s before word-initial stops, nasals, liquids and glides, a subphonemic change that is rarely discussed, even in such detailed treatments of Germanic palatalization as van der Hoek (2010). There are good reasons to argue that nonconforming observations in science should be highlighted rather than ignored or deferred to some future generation of scholars, since it is precisely nonconforming observations that present the greatest challenges to our assumptions about the world. Given the tendency to shunt such cases aside, it seems important to make the following body of data about unusual sound changes in the Austronesian language family available to a wider circle of scholars than areal specialists.

Before I begin let me hasten to add that my initial inspiration for writing this paper did not come from the examination of Austronesian language data, but rather from a routine textbook reconstruction problem involving the Papuan (Eleman Family) languages Toaripi and Orokolo in Campbell (2013: 147–148). I regularly teach Ling. 645, "The Comparative Method", at the University of Hawai'i, and this is the text we have used in two editions for the past several years. Students have found this problem exceptionally difficult, and so did I until I abandoned my prejudices about what kinds of vocalic environments could condition consonant changes. Once this happened everything fell into place, as though the conditioning was as natural as final devoicing, or partial assimilation of *a to the height and frontness of a high vowel in the next syllable.

Briefly, here is the challenge this problem presents: Toaripi and Orokolo have a *t*: *k* correspondence in word-initial position, and in some intervocalic environments. However, they also have a *t*: *l/r* correspondence (where the reader is told to treat *l* and *r* as "the same sound") in some other intervocalic environments. Since *k must be posited to account for a *k*: *k* correspondence it is clear that the *t*: *k* correspondence reflects *t, and it is equally clear that the *t*: *l/r* correspondence also

reflects *t. So, what conditioned the divergent intervocalic reflexes of *t? (Since there is just one reflex word-initially, the problem arises only in intervocalic position.) Tables 1.1 and 1.2 show the relevant examples from Campbell above the dotted lines and additional relevant data from Brown (1986) below the dotted lines. The two tables show the two outcomes of the change of intervocalic *t, and forms that contain two instances of intervocalic *t under different conditioning factors are listed twice, with segments underlined to indicate the relevant correspondence, as with *at̲ute: ak̲ore* in 4 of Table 1.1 and *atut̲e: akor̲e* in 8 of Table 1.2.

Table 1. Examples of medial *t*: *k* and *t: l/r* in Toaripi and Orokolo of Southeast New Guinea

1. *t: k*	**Toaripi /t/**	**Orokolo /k/**	**gloss**
1.	*ete*	*eke*	'vagina'
2.	*tete*	*keke*	'fish scale'
3.	*uti*	*uki*	'bone'
4.	*at̲ute*	*ak̲ore*	'son, boy, male offspring'
5.	*oviti*	*aviki*	'to fetch, bring'
6.	*ivutu*	*avuku*	'sleep, slumber'
7.	*iti*	*iki*	'come from inland'
8.	*oroti*	*iroki*	'canoe'
9.	*maea tete*	*maea keke*	'dwarf, stunted' (*maea* = 'body')
10.	*tetere*	*kekele*	'earth, soil, ground'
11.	*teto*	*keko*	'giant bamboo'
12.	*maea titiri*	*kikili*	'scabies, rash, itch'
13.	*totora-i*	*kokora*	'to sew, stitch'
14.	*tutururu*	*kukururu*	'thundering sound'
15.	*ita ve lat̲i*	*ila ve lak̲i*	'bristles of a pig'
16.	*maeati*	*maeaki lai*	'to praise'
17.	*mati*	*maki*	'string of dogs' teeth (ornament)'
18.	*maea mariti*	*makiri* (< met.)	'shyness, shame'
19.	*mauti*	*mauki*	'thigh'
20.	*metori*	*mekari*	'well, carefully (done)'
21.	*metere*	*mekere*	'silverfish (insect)'
22.	*oti haro*	*oki*	'path, road' (*haro* = 'main, chief')
23.	*ōto*	*oko*	'thatch (of nipa or sago palm)'
24.	*otoare*	*okoare*	'promise, oath, agreement'

Table 1. (*continued*)

1. *t*: *k*	**Toaripi /t/**	**Orokolo /k/**	**gloss**
25.	*opatora*	*opakora*	'frontlet of shell disks'
26.	*ma pete*	*ma peke*	'tranquil, of sea' (*ma* = 'water')
27.	*pitu*	*piku*	'larva of longicorn beetle'
28.	*poti*	*poke*	'long club of palm wood'
2. *t*: *l*/*r*	**Toaripi /t/**	**Orokolo /l/, /r/**	**gloss**
1.	*afutae*	*afurae*	'ashes'
2.	*aite*	*aire*	'after'
3.	*ita*	*ila*	'pig'
4.	*kite*	*kile*	'mat'
5.	*lauta*	*laura*	'Flame tree'
6.	*puta*	*pura*	'cloth'
7.	*uta*	*ura*	'hole'
8.	*atu̲te*	*ako̲re*	'son, boy, male offspring'
9.	*aruita*	*arila*	'kind of snake'
10.	*faiita*	*haira*	'red ochre'
11.	*haute*	*haure*	'abstinence, dislike'
12.	*sito*	*hilo*	'oyster'
13.	*isuta*	*ihura*	'traditional exclamation that embodies the name of the clan *aualare*'
14.	*ita*	*ila*	postposition; 'with, together with'
15.	*ita* 'stone'	*ila*	'axe'
16.	*i̲ta ve lati*	*i̲la ve laki*	'bristles of a pig'
17.	*itaea*	*ilaea*	'sale, deal, purchase'
18.	*itoro*	*iroro*	'proclamation, oration, harangue'
19.	*ivuta*	*ivura*	'iguana'
20.	*karita-i*	*karila*	'to grate, scrape (coconut)'
21.	*saruta*	*karula*	'louse, lice' (*s*: *k* unexplained)
22.	*kite*	*kile*	'mat'
23.	*maea isu̲ta*	*maea ihu̲ra*	'traditional exclamation of triumph'
24.	*mauta itai*	*maura ukai*	'southeast monsoon'
25.	*ōpaita koru*	*opovila kou*	'Morning star'
26.	*vita*	*vila*	'man, male; husband'

Given our usual notions of conditioning, this problem appears opaque. Once I overcame my training in linguistics sufficiently to treat the data without prejudice I assumed that the *t > *k* change in Orokolo happened between vowels of the same height, and *t > *l/r* between vowels of different height. However, this initial generalization was based on just three examples of *t > *-k-*, and although further supporting evidence was found in Brown (1986), a handful of counterexamples was also found, leaving the matter in limbo.

The conditions for *t > *-k-* as against *t > *-l/r-* in Orokolo can now be stated with much greater confidence. Rather than framing *t > *-k-* as happening between vowels of the same height, and *t > *-l/r-* as happening between vowels of different height, it is clear that the first of these change paths was followed between vowels of level or ascending height (*ete*, *oti*, *eto*, etc.), and the second between vowels of descending height (*uta*, *ite*, *ito*, etc.), a condition that is no less surprising than the one I initially assumed. While the data in Campbell provide only 3 examples of the former change and 7 of the latter, the fuller material extracted from the roughly 2,000 base forms and hundreds of additional derivatives in Brown (1986) now provides 28 examples of the former change and 26 of the latter, or 54 in all. Against the 54 cases supporting this very odd condition for a divergent development of earlier *t are 5 possible counter-examples. These are (with expected forms in parentheses):[2]

(1)	*vita-i*	*avika*	'to hang up, hoist up' (**avira)
(2)	*marota*	*maroka*	'pair, set of two' (**marora)
(3)	*popōta*	*popoka*	'a fern: *Dryopteris* sp.' (**popora)
(4)	*itai*	*ukai*	? (**urai)
(5)	*viviuti*	*viviori*	'mountain duck' (**vivioki)

The first of these may be a chance resemblance (Toaripi *-i* is a well-established verbal suffix, but other examples of Orokolo vowel initial trisyllables corresponding to Toaripi disyllables are extremely rare). Likewise, Toaripi *itai* and Orokolo *ukai*, which appear in the expressions *mauta itai* and *maura ukai* 'southeast monsoon' show an irregular correspondence of the initial vowel. Even if all 5 of these nonconforming examples are counted as counterevidence, the stated conditions for the split of *t into *-k-* or *-l/r-* apply in 54 of 59 cases, or over 90% of the time, a clear departure from the null hypothesis.

---

2. Rare examples of a *t*: *t* correspondence are ignored, as in most cases these are marked as products of borrowing from the locally influential Austronesian language Motu (T *pito-pito*: O *pito-pito* 'button', T *pute*: O *pute* 'bag, sack', etc.). In no. 4 for *t > *-k-* I assume *atute, and in no. 20 *metori.

Although its relevance was initially overlooked, there is a second set of intervocalic correspondences that show an apparent unconditioned phonemic split in intervocalic position, this one involving *s*: *h* in some forms, and *s*: *t* in others. Since the first of these correspondences could only reflect *s, it is clear that Proto-Eleman *s became Orokolo *h-, -h-* in some forms, but *-t-* in others. Tables 2.1 and 2.2 show the two outcomes of the change of intervocalic *s. Relevant examples from Campbell are above the dotted lines and further relevant data from Brown (1986) are below the dotted lines.

Table 2. Examples of medial *s: h* and *s: t* in Toaripi and Orokolo of Southeast New Guinea

1. *s: h*	**Toaripi /s/**	**Orokolo /h/**	**gloss**
1.	*ase*	*ahe*	'sugarcane'
2.	*seseroro*	*heheroro*	'thin'
3.	*sisia*	*hihia*	'sour'
4.	*susu*	*huhu*	'plank'
5.	*asa-i*	*aha*	'to grow'
6.	*apisi*	*ahihi*	'midnight'
7.	*avasa*	*avaha*	'lumbar region, lower back'
8.	*easo*	*eaho*	'many-pronged fish spear'
9.	*fasa-i*	*haha*	'to tie up; knot; loop'
10.	*sosori*	*hahari*	'fight, battle, conflict'
11.	*harisu*	*harihu*	'ghost; spirit of the dead'
12.	*sese*	*hehe*	'antennae of a butterfly'
13.	*seseōva-i*	*heheava*	'to examine, test, check'
14.	*seseve*	*hehee*	'maggot, worm'
15.	*sesera*	*hehera*	'a fish: black sole'
16.	*sevese*	*hevehe*	'masked ceremonial cycle'
17.	*posera*	*hohera*	'famine'
18.	*sosoru*	*hohoro*	'firefly'
19.	*posu*	*hohu*	'water pot'
20.	*maea i̱s̱uta*	*maea i̱ẖura*	'traditional exclamation of triumph'
21.	*maso*	*maho*	'spell, charm, magic'
22.	*marase*	*marahe*	'praise'

2. *s: t*	Toaripi /s/	Orokolo /t/	gloss
1.	*saesa*	*haita*	'dish'
2.	*farisa*	*harita*	'arrow'
3.	*taisa*	*kaita*	'paddle'
4.	*marisa*	*marita*	'girl'
5.	*isave*	*aitave*	'oyster, crescent pearl shell'
6.	*isō*	*aito*	'lime stick'
7.	*arakaisa*	*arakaita*	'single log canoe (of Delta)'
8.	*arisa*	*arita*	'coconut shell spoon'
9.	*saesa*	*haita*	'earthware bowl, dish'
10.	*sisa*	*hita*	'final, closing, last'
11.	*sisoro*	*hitolo*	'year'
12.	*isai*	*itau*	'go coastward (?)'
13.	*leisa*	*leita*	'who?'

Surprisingly, this change is also sensitive to the relative heights of vowels surrounding the segment undergoing change, but can be said to operate with opposite effect, since *t lenited between vowels of descending height, while *s lenited between vowels of equal or ascending height. Although there are fewer positive examples of conditioning for this split (35, as against 54 for *t), there are only three known counter-examples, shown here (with expected forms in parentheses):

(6) *asia* *atiha* 'to sneeze' (**ahiha)

(7) *oroasi* *aroate* 'box; pouch, ditty bag' (**aroare)

(8) *misa* *miha* 'k.o. edible crab' (**mita)

The first of these can surely be dismissed, as words for 'sneeze' are onomatopoetic in many languages, and usually contain [s] or [ʧ]. We can thus count 35 positive examples and two negative ones, or a predictability score of over 94%. Since the split of both intervocalic *t and *s is conditioned by the relative heights of flanking vowels, the combined evidence for this odd condition is 89 positive cases against 7 negative ones, for a predictability score of nearly 94%. These results are summarized in (9), where L = level height, A = ascending, and D = descending, and counter-evidence is underlined):

(9)

	*t	*s		
	-k-	**-l/r-**	**-t-**	**-h-**
L	19	1		17
A	9		1	5
D	4	26	13	1

In addition to these intervocalic developments *t regularly became /k/- (never a liquid), and *s regularly became /h/- (never a stop) word-initially. The full range of reflexes is thus *t to /k/ and *s to /h/ word-initially and intervocalically except between vowels of descending height. The fact that both changes are conditioned in the same unexpected way makes any attempt to dismiss the data as a product of chance highly implausible. Moreover, abandoning an appeal to conditioning would force us to accept a substantial number of unconditioned phonemic splits, a solution that most historical linguists would rather avoid if possible.

We are left, then, with a surprising conclusion: speakers of Orokolo (and thus presumably speakers of any language) must have an innate sense of natural classes (in this case vowel height) that can be applied to sound changes, whether these are phonetically motivated or not. In the present case *t > *l* or *r* intervocalically and *s > *h* unconditionally can be seen as ordinary lenitions. For *t > *k* and *s > *t*, on the other hand, a phonetic motivation is less apparent. The first of these changes is known in more than 20 independent cases in the Austronesian language family, but in all of these *k first vacated its position, leaving just two voiceless stops in the system (Blust 2004). By contrast, in Orokolo *t > *k* led to merger of the two phonemes both initially and intervocalically. The Austronesian cases can therefore be seen as structurally motivated (given just /p/ and /t/, the latter was free to appear allophonically at any position from dental to velar), but no such argument can be advanced for Orokolo, which preserved *k as /k/. Since *s and *t did not merge it seems clear that *t lenited to /l/ or /r/ between vowels of descending height before *s underwent fortition to *t* in the same environment. Again, however, it is not easy to see a structural motivation for this change, since it had no impact on the constitution of the phoneme system. Simple learning of patterns and generalizing them to other cases seems inadequate to account for the intuitive sense of vowel height differences in these two sets of changes, since in the published data there are three separate patterns for *s > /t/ (*esa, isa, iso*), five for *t > /l/ or /r/ (*ita, ite, ito, uta, ite*), 10 for *s > /h/ (*asa, ase, aso, ese, ose, oso, osu, isi, isu, usu)*, and 11 for *t > /k/ *(ati, ato, atu, ete, eto, oti, oto, iti, itu, uti, utu*).

3. Odd conditioning of vowel changes

The following six case studies focus on vowel changes that may be natural in themselves, but which are conditioned in ways that no current theory of phonology would lead us to expect. The first is the rounding or fronting of word-final low vowels in a number of Austronesian languages, primarily in insular Southeast Asia, but reaching from Taiwan to the Solomon Islands. The second is the fronting of *u in the final syllable (whether open or closed) in a number of the languages of eastern Indonesia. The third is the raising of *a adjacent to the uvular stop *q in Kavalan of northeast Taiwan. The fourth is a process of vowel breaking in languages of coastal Sarawak, with some apparent diffusion of this change into inland communities. In most languages, vowel breaking involves the development of a mid-central offglide to high vowels before word-final *k and *ŋ (but not *g) and of a mid-central onglide to word-final high vowels. In the Mukah dialect of Melanau, breaking was extended to *a, which was raised to *e* and offglided exactly like the high vowels. The fifth, which is found in the recently-described Sula language of eastern Indonesia, is a lowering of *u (but not *i) in the penult if the following syllable contained a high vowel (*i or *u). In the Sanana dialect *-u then became *a* if the penult contained *o*. The last case study involving vowels is low vowel dissimilation, a recurrent change throughout the Oceanic branch of Austronesian, which occupies most of the Pacific islands. In this change, the first of two low vowels was raised to *e* or *i*, leaving a synchronic residue in some languages in which the unaffixed base has a single vowel *a*, and the affixed base has eCa or iCa < *aCa.

3.1 Case study 1: Final /a/ mutation

With over 90 million speakers, Javanese is easily the largest language in the Austronesian (AN) family, and one of its better-studied members as a result of a long tradition of Dutch philological research starting in the first half of the 19th century (Uhlenbeck 1964), followed by increasing international attention after Indonesia achieved its independence in 1949.

It is well known that Proto-Austronesian *a became Javanese *ɔ* word-finally, and that a low vowel in a preceding syllable assimilated completely to it, a process that is blocked by suffixation, as seen in *lima > *limɔ* 'five', or *mata > *mɔtɔ* 'eye', next to e.g. *lima-ŋ* 'five (as modifier)', *kə-mata-n* 'having excessively large eyes' (Dempwolff 1934: 29, Horne 1974: xiii–xiv). Several other languages in western Indonesia show different changes that are also restricted to word-final *a, as with the fronting and raising of this vowel in Jakarta Malay (*lima > *limɛ* 'five', *mata > *matɛ* 'eye'), or the raising of word-final *a to schwa in Balinese and some nonstandard dialects of Malay. Tadmor (2003) has coined the useful term "final /a/

mutation" for this phenomenon, and I follow his usage here. Tadmor's explanation for final /a/ mutation takes essentially the following form:

- Many Malay dialects have raised *-a to *-ə*, *-o* or *-e*.
- A similar change is found in Lampung, Javanese, Balinese and other languages of western Indonesia. It can thus be characterized as an areal feature, that is, a by-product of language contact and bilingualism.
- But how would a change that has no clear phonetic motivation begin? To solve this problem, Tadmor (2003: 21) posits final /a/ mutation in the Javanese-speaking Hindu-Buddhist state of Majapahit (1293–1500), where Indian Brahmans introduced large numbers of Sanskrit loanwords at a time when short /ă/ had begun to raise. Because they were inundated with loans that had this phonetic feature, speakers of Old Javanese, the court language of east Java from the 9th to the 15th century, applied final /a/ mutation to their native vocabulary.
- The conquests of Majapahit, which created a medieval empire encompassing much of what is now Indonesia-Malaysia, led to Javanese influence on the languages of subject peoples, who borrowed final /a/ mutation as a prestige feature, much as the uvular /r/ of Parisian French was borrowed into other languages of Europe.
- What was borrowed was a centralized reflex of *-a (presumably [ə]). Because the languages that adopted this feature reportedly have iambic stress "the various 'full-vowel' reflexes of *-a emerged as a restrengthening of a centralised vowel of some sort" (Tadmor 2003: 31).

Tadmor concludes by acknowledging that final /a/ mutation (*a > *-o* in some languages, *a > *-e* or *-ɛ* in others) occurs outside the area where Javanese contact influence is attested, but he argues that this fact does not compromise his claim that the feature was spread through western Indonesia by speakers of Javanese during the apogee of Majapahit military and political power. Again, he appeals to the analogy of uvular /r/ in Europe: just because some languages outside the areal influence of Parisian French have changed an apical to a uvular /r/ does not mean that others did not adopt this feature through its perceived prestige value when French was the preferred language of culture and diplomacy in Europe.

This is an interesting argument, but one that raises more questions about the motivation of sound change than it answers. First, Old Javanese had a schwa that was found in most positions, but not word-finally, and it remained a schwa in all known dialects of modern Javanese. To claim that *-a became mid-central and then rounded simply shifts the onus of explanation from *a > *-ɔ* to *ə > *-ɔ*. Moreover, the proposed impetus for the change – that /ă/ in Sanskrit was raised and centralized, and that this feature was then transferred to native Javanese vocabulary – is based

on little more than speculation. As Tadmor notes, the centralization of /a/ in some descendants of Sanskrit, as Hindi, has affected short /a/ in all positions, not just word-finally. To the extent that the phonetics of the modern Indic languages can be taken as a guide to the phonetics of Sanskrit, then, it must be asked why speakers of Old Javanese would have copied the raising and centralization of Sanskrit short /a/ only where it occurred in an open final syllable. The answer may be that phonetic features are more easily copied where they lead to shift than to merger: /a/ and /ə/ contrasted in Old Javanese, but word-finally only *-a* was found, leaving open the possibility for phonetic manipulation without risking the adoption of a phonetic feature than could potentially have led to increased homophony.

Even if we acknowledge this possibility, Tadmor assumes that all instances of final /a/ mutation passed through a stage in which the final vowel was mid-central, not just in Javanese, but also in other languages of western Indonesia, and that the subsequent acquisition of labiality or palatality was conditioned by stress. As will be seen below, the first part of this assumption can be justified for languages that have an identical reflex for *-ə and *-a, but it is far harder to maintain for languages that do not. In modern standard Javanese, for example, the schwa remains mid-central. As already noted, if *-a > [ɔ] passed through a stage *-a > [ə] the language then underwent *ə > [ɔ] only word-finally. The claim that this hypothetical change was stress-conditioned conflicts with two fundamental observations: first, stress in Javanese is not (and, so far as we know, never was) sensitive to syllable shape, and second there is no phonetic evidence that stress would make schwa more prone to the acquisition of labiality or palatality (which is obviously distinct from the *loss* of labiality or palatality in unstressed position).

When we look at the distribution of such changes in a wider sample of languages it is difficult to escape the conclusion (pace Tadmor) that final /a/ mutation, as a general pattern of phonological innovation, is neither borrowed nor an areal phenomenon. Rather, in many cases it is a system-internal change that lacks any obvious phonetic motivation. To see this it is necessary to briefly survey the historical phonology of a representative sample of other languages, many of which are spoken by traditional animists who had little if any contact with the Indianized states of insular Southeast Asia.

Table 3 summarizes the reflexes of Proto-Austronesian (PAN) final *-a, non-final *a, and schwa in Austronesian languages reaching from south-central Taiwan to the Solomon Islands ($ə_1/o_2$: = /ə/ in the penult, /o/ in the ultima).

Several patterns are observable in this table. First, where *-a has the same reflex as *-ə, as in Central Cagayan Agta (2) or Tombonuwo (6), it is more difficult to argue that it rounded without passing through an intermediate stage in which it merged with schwa. This eliminates all cases except 3, 8–10, 12, 13, 16, 18–22, 26, and possibly 17 and 25. Second, where both word-final *a and non-final *a

Table 3. Candidates for final /a/ mutation

	PAN	***-a**	***a**	***ə**
1.	Tsou (Taiwan)	-o	a/o	ə
2.	Central Cagayan Agta (N. Philippines)	-ʌ	a/ʌ	ʌ
3.	Tiruray (S. Philippines)	-o	o	ə
4.	Koronadal Bilaan (S. Philippines)	-ɨ/ɔ	a/ɨ/ɔ	a/ɨ
5.	Ida'an Begak (E. Sabah)	-o	a	$ə_1/o_2$
6.	Tombonuwo (E. Sabah)	-o	a	o
7.	Kadazan (W. Sabah)	-o	a	o
8.	Timugon Murut (W. Sabah)	-o	a	a
9.	Kedayan (Brunei)	-o	a	a
10.	Bisaya Bukid (Brunei)	-o	a	ə
11.	Tutong (Brunei)	-o	a	$ə_1/o_2$
12.	Long Dunin Kenyah (Sarawak)	-e	a	ə
13.	Òma Lóngh Kenyah (C. Kalimantan)	-o	a/ɛ	ə
14.	Bekatan (S. Sarawak)	-o	a	o
15.	Maanyan (S.E. Kalimantan)	-e	a	e
16.	Malagasy (Madagascar)	-i	a	e
17.	Simalur (N. Sumatra)	-a/o	a	ə/a
18.	Minangkabau (C. Sumatra)	-o	a	a
19.	Kerinci (C. Sumatra)	-o	a	ə
20.	Lampung (S. Sumatra)	-o	a	ə
21.	Jakarta Malay/Betawi (W. Java)	-ɛ	a	ə
22.	Javanese (Java)	-ɔ	a	ə
23.	Gorontalo (N. Sulawesi)	-o	a	o
24.	Bolaang Mongondow (N. Sulawesi)	-o	a	o
25.	Kédang (Lesser Sundas)	-e/o	a	è/e
26.	Paripao (Guadalcanal, Solomons)	-e/a	a	o

have the same reflex, as in Tsou (1) or Tiruray (3), it is impossible to distinguish final /a/ mutation from general rounding of *a. Third, some of these languages are in the "greater Indic diffusion sphere" that dominated Javanese, Malay, and some other languages of western Indonesia for several centuries, and so are susceptible to the alternative explanation that Tadmor proposes. This includes 18–22, but not 8–10, 12, 13, 16, 25, 26, or possibly 17. It will be useful to look first at Long Dunin and Òma Lóngh, two distinctly different dialects of Kenyah spoken

in central Borneo that had little contact with the outside world until the late 19th or early 20th century.

Table 4 shows reflexes of Proto-Malayo-Polynesian (PMP, the immediate ancestor of all non-Formosan Austronesian languages) and Proto-North Sarawak (PNS, the immediate ancestor of 15–20 languages in northern Sarawak) in Long Dunin, a longhouse located on the right bank of the Tinjar river in northern Sarawak some 50 miles upriver from its junction with the Baram, in an area that had little or no contact with Malay until the past several generations.

Table 4. Evidence for *-a > -e ([ɛ]) in Long Dunin Kenyah

1. PMP and Long Dunin		
	PMP	**Long Dunin**
	*dəpa	*dəpe* 'fathom'
	*duha (> *dua)	*due* 'two'
	*ia	*ie* '3.SG'
	*lima	*ləme*[a] 'five'
	*laqia (> *lia)	*lie* 'ginger'
	*mata	*mate* 'eye'
	*tawa	*tawe* 'to laugh'
	*taliŋa	*təliŋe ~ kəliŋe* 'ear'
	*tuba	*tuve* 'fish poison: *Derris elliptica*'
	*quma	*ume* 'swidden'
2. PNS and Long Dunin		
	PNS	**Long Dunin**
	*əj^{h}a	*se* 'one'
	*usa	*use* 'body'

a PMP *lima > *ləma was an irregular change in Proto-Kenyah.

As can be seen, *a and *ə normally remained unchanged, except that word-final *a was fronted to [ɛ]. A similar change did not occur to last-syllable *a if it is now, or was earlier, followed by a consonant, as in *anak > *anak* 'child', *buaq > *buaʔ* 'fruit', *daRaq > *daaʔ* 'blood', *kulat > *kulat* 'mushroom', *ŋajan > *ŋaran* 'name', *qudaŋ > *uraŋ* 'shrimp', or *dipaR > *kə-dipa* 'opposite bank', *beRas > *baa* 'husked rice', *panas > *pana* 'warm, hot', *teRas > *taa* 'ironwood tree'. Note also that PMP *ə remained a mid-central vowel except word-initially, where it disappeared: *təlu > *təlәw* 'three', *ənəm > *nəm* 'six', *quləj > *ulət* 'maggot', *daləm > *daləm* 'inside', *ləñəb > *ləñəp* 'flood', *bəsuR > *bəso* 'satiated', etc. Likewise, word-final /a/ in Malay loanwords

has not been changed: *basa* 'read' (< Malay *baca*), *buka > *buka* 'open' (< Malay *buka*), *buŋa > *buŋa* 'household flower' (< Malay *buŋa*), etc. In conclusion, final /a/ mutation in Long Dunin Kenyah shows a remarkable parallelism with Jakarta Malay/Betawi. In both languages *-a became [ɛ], a sound that is phonemic in Jakarta Malay (because of longstanding influence from Javanese, which contrasts /e/ and /ɛ/, /o/ and /ɔ/), but remains an allophone of /e/ in Long Dunin Kenyah.

The historical phonology of Òma Lóngh Kenyah is treated at some length in Blust (2007), and the relevant portions will be summarized here only briefly. This language has undergone extensive sound changes which are often conditioned in ways that complicate the statement of how a final low vowel developed. However, the usual path was *-a > *-o*, much as in Javanese, except that /o/ is [o] rather than [ɔ] (Soriente 2006: xl–xli). Representative etymologies are shown in Table 5.

Table 5. Evidence for *-a > *-o* ([o]) in Òma Lóngh Kenyah

1. PMP and Òma Lóngh		
	PMP	**Òma Lóngh**
	*duha	*dəvo/rəvo* 'two'
	*ia	*zo* '3.SG actor'
	*lima	*ləmo* 'five'
	*laqia (> *lia)	*ləzo* 'ginger'
	*mata	*mato* 'eye'
	*mula	*molo* 'to plant'
	*-ni-á (> *-ña)	*-ño* '3.SG possessor'
	*daya	*razo* 'upstream, inland'
	*tuqa	*toʔo* 'elder; parents'
	*tawa	*tavo* 'to laugh'
	*taliŋa	*təliŋo* 'ear'
	*tuba	*tubo* 'derris root fish poison'
	*uma	*omo* 'swidden'
2. PNS and Òma Lóngh		
	PNS	**Òma Lóngh**
	*əj^{h}a	co 'one'
	*usa	oso 'body'

Òma Lóngh is spoken on the Malinau river in a remote region of East Kalimantan, and like Long Dunin, had little contact influence with Malays or other non-Bornean peoples until relatively recent times. Contact influence as a cause of final /a/ mutation is thus extremely unlikely.

Finally, we can consider Kédang, a Central Malayo-Polynesian language spoken by a culturally conservative people located on the small island of Lembata in the Solor archipelago between Flores and Alor, eastern Indonesia. Kédang is surrounded by languages in the extensive Lamaholot dialect chain, but stands apart from them (Barnes 1974: 19–20, Fox n.d.). Keraf (1978), who treats Kédang as part of this continuum, provides a thumbnail sketch of variation within Lamaholot based on a sample of 35 distinct communities, and apart from Kédang only one of these (Mulan) shows any tendency to final /a/ mutation.

The detailed ethnography of Barnes (1974) suggests that the Kédang people experienced little if any cultural or linguistic influence from western Indonesia until very recent times. In particular, as in the case of Long Dunin and Òma Lóngh Kenyah, the likelihood of Javanese influence on the phonological development of this language, is virtually nil. Final /a/ mutation in Kédang is especially challenging, since it appears to involve an unconditioned phonemic split, the final low vowel becoming either /e/ or /o/ with roughly equal frequency, and without statable conditions. Evidence for this change, based on an unpublished vocabulary of about 1,200 words supplied by the Dutch anthropologist Martine van Dierendonck (van Dierendonck n.d.), appears in Table 6.

Table 6. Evidence for final /a/ mutation in Kédang

PMP	**Kédang**
*ama	*amo* 'father'
*qabaRa	*bare* 'shoulder; carry on shoulder'
*apa	*ape* 'what?'
*daya	*hoyo* 'upriver'
*ina	*ino* 'mother'
*lima	*leme* 'five'
*laqia (> *lia)	*lie* 'ginger'
*mata	*mato* 'eye'
*pija	*pie* 'how much/many?'
*duha	*sue* 'two'
*i sai (> *i sia)	*sio* 'who?'
*tawa	*tawe* 'to laugh'
*kita	*te* 'we (incl.)'
*tumah (> *tuma)	*tumoʔ* 'clothes louse'
*tuqah	*tuwo lahar* 'elder; elders'
*buaq (> *bua)	*ue* 'areca nut'

According to Barnes (1974: 20), "Disregarding slight variations of pronunciation and vocabulary, the language is homogeneous throughout Kédang". Several of the forms in the 200-item Swadesh list of Keraf (1978) nonetheless differ from those in my primary source, and these provide both additional examples of final /a/ mutation, and evidence for *-o* and *-e* variation within the language. Examples of the first type include *um-utaq > muta > *mute* 'vomit' (van Dierendonck: *okaʔ*) and *betaq (> beta) > *bete* 'split' (no equivalent in van Dierendonck). Examples of the second type include *ame* 'father', *ine* 'mother' (van Dierendonck: *amo*, *ino*).[3]

In non-final position *a did not normally change in Kédang, as seen in several of the above examples, and in additional etymologies such as PMP *kahiw > *ai* 'wood, tree', *qañam 'plait' > *anaŋ* 'weave', *anak > *anaʔ* 'child', *haŋin > >*aŋin* 'wind', *hikan > *iʔa* 'fish', *zalan > *lala* 'path, road', *laki > *laʔi* 'male', *liaŋ > *liaŋ* 'cave', *manuk > *manuʔ* 'chicken', *nanaq > *nana-n* 'pus', *ŋajan > *naya* 'name', *panas > *pana* 'hot', *Ramut > *ramuʔ* 'root', *qudaŋ > *uraŋ* 'shrimp', *buaq > *ua-n* 'fruit', *quzan > >*uya* 'rain', or *wanan > *wana* 'right side'.[4]

Kédang, then, presents a complex picture in which a number of forms reflect final *a as a mid-front vowel which may have passed through an intermediate stage with schwa, but a number of others reflect final *a as *o*, which evidently involved a direct transition from *a. The latter forms thus show a change that is strikingly similar to that in Javanese and Òma Lóngh Kenyah, although it is regular in the latter languages, but not in Kédang.

What can we make of this wider picture of sound change in the Austronesian family? First, final /a/ mutation clearly is not confined to languages that are likely to have acquired it through contact from Javanese. Tadmor anticipated this objection with regard to Gorontalo in northern Sulawesi, out of the reach of medieval Javanese cultural or political influence (Table 3), and he argues that this fact does not compromise his claim that the feature was spread through western Indonesia by speakers of Javanese during the apogee of Majapahit military and political power. As noted already, in support of this view he appeals to the spread of uvular /r/ in

---

3. Fox (n.d.) gives *ame*/*amo* 'father', and the possessed form *ina-g* 'my mother', which implies that /a/ alternates with /e/ or /o/ in suffixed bases.

4. Van Dierendonck writes the reflex of PMP *ə as *e*: *dəŋəR > *deŋer* 'to hear', *ənəm > *eneŋ* 'six', *pənuq > *penu* 'full', *qapəju > *peyu-n* 'gall', *pusəj > *puhe* 'navel', *sələm > *heleŋ* 'to set, of the sun', *tanəm > *taneŋ* 'to bury', etc. This corresponds (at least orthographically) with one of the reflexes of *-a, but not to the other. Samely (1991), on the other hand, distinguishes *è* ([æ]), from e ([ɛ]) , as in *dəŋəR > *dèŋèr* 'hear, listen', *əpu > *èpu* 'grandparent', *kəmuR > *èmur* 'gargle', *ənəm > *ènèŋ* 'six', *pənuq > *pènu* 'full', *qapəju > *pèyu-n* 'bile, gall', *təbuh > *tèhu* 'sugarcane', or *təlu > *tèlu* 'three', next to *qənay 'sand' > *ene* 'fine sand', *pukət > *pue* 'dragnet', *pusəj 'navel' > *puhe* 'center, middle', *bəli 'buy' > *weri-n* 'bridewealth', or *baqəRu (> *bəRu) > *weru* 'new'. However, this has no bearing on *-a > *o*.

Europe: just because some languages outside the areal influence of Parisian French have changed an apical to a uvular /r/ does not mean that others did not adopt this feature through its perceived prestige value. However, it seems to me that this somewhat defensive posture misses the essential point of what makes final /a/ mutation theoretically intriguing.

Tadmor's approach to this problem clearly began with the recognition that final /a/ mutation lacks self-evident phonetic motivation, but in the course of his argument he becomes more concerned with the issue whether *a > *ɨ*, *ə*, *e* or *o* is a phonetically natural change than with the more puzzling issue of why such a change, if it occurred at all, would occur only *word-finally*. There is, of course, nothing surprising in itself about a conditioned change *a > ɔ, *a > *o* or *a > ɛ. However, as linguists we are accustomed to expect such changes to occur in environments in which they reflect the influence of neighboring vowels (or, more rarely, consonants): *a > ɔ or *o*/__Cu, *a > ɛ or *e*/__Ci, etc. What is puzzling about final /a/ mutation, and particularly about its recurrent nature across a number of languages, is the absence of any obvious reason why word-final position should be a condition for either rounding or fronting. To compound the enigma, as Donegan (1993) points out with respect to the relative stability of vowels in historical change:

> We are left with what Hellwag called the Prince of Vowels, [*a*], which would seem to present no difficulties at all: it is optimally open and has no conflicting features like palatality or labiality. If a child can say only one vowel, that vowel is [*a*]. Or if a language has only one vowel, we would expect it to be [*a*]. And we might expect that [*a*] would always remain [*a*]. But we find that [*a*] sometimes does undergo substitution, either acquiring palatality ([*a*] → [æ]) or labiality ([*a*] → [ɒ]), or losing some of its sonority ([*a*] → [ʌ]). Donegan (1993: 101)

Donegan goes on to note the kinds of conditions that commonly favor the acquisition of palatality or labiality, or the loss of sonority, but obviously word-final position is not among them.

In addition, it might be noted that areal features in both language and culture are products of intense contact, and tend to blanket contiguous geographical areas, not to produce a scattered hit-and-miss distribution such as we find with this pattern of sound change. Linguists hardly need to be reminded of the geographical compactness of areal features such as the use of retroflex stops in India, the absence of nasal consonants in various genetically diverse languages of the Northwest Coast, or the syncretism of the genitive and dative cases in the Balkan Sprachbund. It is true that linguistic areas may show gaps; as noted by Comrie (1989), for example, Serbo-Croatian, which otherwise participates in the Balkan Sprachbund, distinguishes genitive from dative case, but it is clearly inappropriate to use the term 'areal feature' for a phenomenon that is scattered over a wide area

with many intervening non-conforming cases. In the present case final /a/ mutation, as reflected in Table 3, is attested from south-central Taiwan to the central Solomon Islands, hardly what anyone would seriously call a "linguistic area".

Finally, even in the paradigm case of Javanese, as noted by Nothofer (1981), the rounding of low vowels in word-final position took place in some Javanese dialects or villages (including the standard language), but not in others. Since the kingdom of Madjapahit was centered in East Java, it would be surprising indeed if the influence of the most prestigious dialect of Javanese had penetrated so many foreign linguistic zones while leaving significant parts of the home territory untouched.

In short, Tadmor deserves credit for noting the distribution of final /a/ mutation in some of the better-known languages of western Indonesia, and for creating a label that facilitates reference to it. However, his claim that it is a borrowed areal feature fails to carry conviction, first for the obvious reason that it occurs in languages that have never been subject to significant Javanese contact influence, and second because it lacks the commonly accepted character of an areal feature.

3.2 Case study 2: Fronting of last syllable *u

Final /a/ mutation is undoubtedly one of the more puzzling conditioned vowel changes in Austronesian languages, but it is not alone. Perhaps equally puzzling is *u > /i/ only in the final syllable, a change found in a number of the languages of eastern Indonesia. These include Wetan of the Babar archipelago in the southern Moluccas, Bobot (Bonfia) of the central Moluccas, and several (but not all) languages of the South Halmahera-West New Guinea group in the northern Moluccas and the Bird's Head Peninsula of western New Guinea, here represented by Buli of southern Halmahera and Waropen and Munggui of western New Guinea. Reflexes of PMP forms from all three areas are given in (10):

(10) a. Wetan (de Josselin de Jong 1987): *asu > *ai* 'dog', *inum > *emni* (met.) 'drink', *kutu > *kuti* 'louse', *tunu > *tuni* 'to burn', *tuktuk (> *tutuk) > *tuti* 'knock, bump', *susu > *ui* 'female breast', *suluq > *uli* 'torch', *qulu 'head; first, first-born' > *uli* 'first, formerly; in front', *batu > *wati* 'stone', *təlu > *wo-teli* 'three', *bulu 'body hair; feathers' > *wuli* 'hair', *bunuq > *wuni* 'to kill', *bubu > *wuwi* 'fish trap'.

b. Bobot/Werinama (Greenhill, Blust & Gray, 2003–2017; also known as Bonfia, although the latter may be a dialect): *təlu > *ain-toli* 'three', *batu > *hati* 'stone', *bunuq > *huin* (met.) 'to kill', *bulu > *hul-huli-n* 'feather', *takut > *ka-toit* 'to fear', *inum > *n-imin* (met.) 'to drink', *susu > *susi-n* 'female breast', *qatəluR > *toli-n* 'egg', *tunu > *tuin* (met.) 'to burn', *qulu > *uli-n* 'head'.

c. Buli: PMP *ma-takut > *am-cait* 'afraid' (cp. *pa-takut > *fa-cait* 'to frighten'), *ma-tuduR > *am-tuli* 'asleep', *hipun > *ifiŋ* 'tiny fish that ascends rivers in huge swarms to spawn', *lusuŋ > *lusiŋ* 'rice mortar', *manuk > *mani* 'bird', *busuR > *pusi* 'bow; shoot with bow', *timuR 'southeast monsoon' > *simi* 'south, south wind', *timun > *ti-timin* 'cucumber', *Rusuk > *usi* 'ribs', *walu > *wai* 'eight' (*-l- > Ø unexplained).

In each of these languages the change *u > /i/ has led to merger with *i: PMP *paRih > Wetan *ari* 'stingray', *lilin > *lilli* (met.) 'beeswax', *tali > *tali* 'rope', *buRiq > *wuri* 'to wash, as the hands', PMP *ma-putiq > Bobot *buti* 'white', *piliq > *pili* 'to choose', *tasik > *tasiʔ* 'sea, saltwater', PMP *paniki > Buli *fni* 'flying fox', *laŋit > *laŋit* 'sky', *mamin > *mamiŋ* 'wrasse', etc.

Buli has some unexplained exceptions, as with *lahud > *lau* 'seaward', *zauq > *lau* 'far', *bulu > *plu* 'body hair, feathers', *taqun > *taun* 'year' and *qutub > *utup* 'submerge a container to fill it'. However, in most cases, last-syllable *u became *i*. Because Buli lost original final vowels, a number of potential examples of this change are no longer available. This change also occurs sporadically in several languages of the West New Guinea group, but apparently was not present in their immediate ancestor: PMP *qasu > Waropen *asi* 'smoke; steam', *qulu > *uri* 'upstream', *kubi* (< Malay *kubur*) 'grave', *manuk > Waropen *mani* (but Moor *manu*) 'bird, chicken', *susu > Munggui, Waropen *susi* (but Pom *huhu*, Wandamen *susu*) 'female breast', etc.

A comparison of this conditioned vowel change with final /a/ mutation raises some interesting questions. Both changes affect a vowel only in the final syllable: final /a/ mutation only if it is open, and *u > /i/ whether it is open or closed. In neither case is there a discernible connection between the nature of the change and the phonetic properties of the conditioning environment. The most puzzling feature of both changes is their recurrence across a population of languages that do not form an immediate subgroup, since this distribution implies a shared motivation. In other words, each change has the form of a drift, and so implies the continued operation of an inherited structural pressure after the separation of the daughter languages from a common ancestor. But what this structural pressure might have been remains a mystery, hence Tadmor's appeal to contact for final /a/ mutation, as a kind of "escape" from the prima facie contradiction of drift without motivation. Since there is no structural condition favoring *u > *i*/__(C)# in languages of eastern Indonesia but not elsewhere in the Austronesian family, the cross-linguistic distribution of the latter change is doubly puzzling: it patterns like a drift, but only in a restricted geographical area where it does not appear to be due to contact.

Perhaps because of interaction with stress, in many Austronesian languages the inherited schwa has different reflexes in the penult and ultima. Moreover, *ə has merged with each of the other vowels in one language or another, a pattern that is unsurprising, since a mid-central vowel is not distant in articulatory space from any other vowel. However, it is unusual for the maximally distinct high vowels to merge, and there is no obvious reason why it would be more likely for this change to happen in the ultima rather than the penult (examples of *u > i only in the penult are unknown in Austronesian languages). To conclude, we are left with the seemingly unanswerable question "What is it about the final syllable of a word that favors vowel fronting?".[5]

3.3 Case study 3: *a > i adjacent *q in Kavalan

PAN *q was a uvular stop, and although it became *ʔ*, *k*, *h*, or zero in nearly all languages outside Taiwan, it is reflected unchanged in a number of the Formosan languages. As in Arabic or Quechua, [q] lowers adjacent high vowels, altering /qu/ to [qo], [oq], and /qi/ to [qəj], [eəq], a development illustrated by the Thao (central Taiwan) words /qusað/ [qósað] 'rain', /punuq/ [púnoq] 'head', qiła/ [qə́jɬa] 'rice wine', or /apiq/ [ápeəq] 'daughter-in-law'.

In Kavalan of northeast Taiwan (Li & Tsuchida 2006), *q disappeared but left a trace on adjacent vowels. However, the trace was not the expected lowering effect on high vowels, but rather a raising effect on the low vowel *a, as seen in Table 7.

In four known examples *a > *i* when not adjacent to *q. Three of these are enclitic pronouns, where *a > *i* in initial position: *aku > *-iku* '1.SG', *ami > *-imi* '1.PL ex', and *amu > *-imu* '2.PL'. The fourth shows word-final *a > *i*: *depa > *zəpi* 'fathom'. The explanation for these cases is unclear. Otherwise, as noted in Li & Tsuchida (2006), adjacency to *q is the sole condition under which *a raises (cf. *ajem > *anem* 'heart, mind', *ala > *ara* 'get, fetch', *lima > *rima* 'five', *t-ina > *tina* 'mother', *daya 'inland' > *zaya* 'west', *quzaN > *uzan* 'rain', bulaN > *buran* 'moon, month', *Sawak > *sawaq* 'back of the waist', *huRaC > *uRat* 'blood vessel', *damuq > *zamu* 'menstruation'), and there is no lowering of high vowels in *qiNeb > *ineb* 'door', *qipit > *ipit* 'tongs', *qumaŋ > *umaŋ* 'hermit crab', *qulu > *uRu* 'head', *quay > *uway* 'rattan', etc.).

Nearly all Formosan languages preserve PAN *k as a velar stop, but in Kavalan after low vowels were raised the *q was lost, and *k became *q* under conditions that

5. The merger of *u and *i has also occurred in a number of Austronesian languages as a *sporadic* change (Blust 1970), but in this case, it does not favor a particular syllable. It is nonetheless noteworthy that whenever the contrast of *u and *i is neutralized by merger the favored outcome is [i], not [u].

Table 7. Raising and fronting of *a next to *q in Kavalan

PAN	Kavalan
*qabu	*ibu* 'ash'
*qamiS	*imis* 'north'
*qalad	*iRaR* 'gate; fence'
*qaRaw	*iRaw* 'snatch, seize, rob'
*qaRem	*iRem* 'pangolin'
*qaluR	*iRuR* 'small stream'
*qatəb	*itəb* 'deadfall trap'
*qauR	*iuR* 'slender bamboo'
*nanaq	*ku-nani* 'pus'
*ma-qetaq	*məti* 'raw, unripe'
*panaq	*p<m>ani* 'shoot with arrow'
*luSəq (> *luSaq)	*Rusi* 'tears'
*Samaq	*sami* 'common sow thistle'
*tanaq	*tani* '*Aralia decaisneana*'
*taRaq	*t<m>ari* 'chop with an axe'
*utaq	*uti* 'vomit'

are not fully statable.[6] What is noteworthy, and a further cause of puzzlement, is that /a/ now occurs next to /q/ from *k without the raising effect caused by PAN *q, as seen in e.g. *kaka > *qaqa* 'elder sibling (ref.)', *kaRaC > *qaRat* 'to bite', *baki > *baqi* 'grandfather', *pukaw > *puqaw* 'blind', or *pakpak > *p<m>aqpaq* 'flap the wings'. One would like to attribute this difference in the behavior of [a] to a difference in the phonetic properties of the original and secondary phonemes written *q*, but Kavalan /q/ is a voiceless uvular stop, and it is universally assumed that PAN *q was as well. Given this conclusion, and the unnatural process of raising low vowels adjacent to a uvular stop, there appears to be no basis for arguing that the vocalic changes seen in Table 7 are phonetically motivated. Rather, Kavalan speakers appear to have had deliberate control over raising *a adjacent to *q in the first instance, and simply did not invoke this option in the second instance, when a new [q] arose from *k.

6. Li & Tsuchida (2006: 7) say that *k > *q*/*u,*a, and *k > *k* elsewhere, but they note exceptions that compromise this statement.

3.4 Case study 4: Vowel breaking in coastal Sarawak

A number of coastal and near-coastal languages of northern and central Sarawak have undergone a phonetic process of vowel breaking that differs in detail from one community to the next, but shares certain common features. This includes languages in the extensive Melanau dialect chain that stretches for about 150 miles along the coast of Sarawak from the mouth of the Rejang River in the south to Balingian in the north, languages in the Lower Baram branch of the North Sarawak subgroup, such as Miri, Narum, and Kiput, and some languages from other groups, such as the Uma Juman dialect of Kayan, which differs from nearly all other known Kayan dialects in possessing this feature (Blust 1977).

Vowel breaking is perhaps best known from the history of the Romance languages, where Latin short *ĕ* and *ŏ* in stressed syllables were diphthongized in Proto-Romance (Posner 1996: 157-). The Austronesian situation is quite different, as in many cases the vowel that diphthongized is unstressed, diphthongization shows different patterns of breaking in open and closed syllables, both high and low vowels were targeted, the latter only with raising, and there are major cross-linguistic differences in the class of final consonants that acted as a conditioning factor, suggesting an areal feature that was reinterpreted by each new community that adopted it.

Table 8 provides an overview from a sample of ten language communities that exhibit at least some vowel breaking. The leftmost column shows the reconstructed endings that underwent breaking in one or more languages. Where a change that appears to be related to breaking occurs, it is indicated in the appropriate column by phonetic transcriptions of the outcome. These transcriptions include breaking, and the lowering of high vowels and raising of low vowels in the environment in which breaking occurs in other languages. Where no relevant change has occurred, the space is left blank.

Table 8. Vowel breaking in ten languages of coastal Sarawak[a]

	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**
*-ik	[iək]	[iəʔ]	[eək]	[iək]	[eək]	[iəʔ]	[iək]	[ek]	[aik]	[iəʔ]
*-iŋ	[iəŋ]	[iəŋ]	[eəŋ]	[iəŋ]	[eəŋ]	[iəŋ]	[iəŋ]	[eŋ]	[aiŋ]	[iə]
*-uk	[ok]	[uəʔ]	[oʔ]	[uək]	[ok]	[uəʔ]		[ok]	[auk]	[uəʔ]
*-uŋ	[oŋ]	[uəŋ]	[oŋ]	[uəŋ]	[oŋ]	[uəŋ]		[oŋ]	[uəŋ]	[uə]
*-it		[iət]						[ait]		
*-ut		[uət]						[ut]	[aut]	
*-iʔ		[iəʔ]		[əiʔ]		[eʔ]		[aiʔ]		
*-uʔ		[uəʔ]		[əuʔ]		[oʔ]		[auʔ]		

Table 8. (*continued*)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
*-in								[ain]		
*-un		[uən]						[on]	[aun]	
*-ul		[uəl]								
*-iR		[iəh]		[əih]	[eγ]				[aiγ]	
*-uR		[uəh]		[əuh]	[oh]					
*-aR						[eəh]				
*-is		[is]								
*-us		[uəs]								
*-ak			[iʔ]			[eəʔ]			[ak]	
*-aŋ			[i]			[eəŋ]			[iəŋ]	
*-an									[iən]	
*-i	[eʔ]	[aj]	[əj]	[əj]	[əj]	[əj]	[əj]	[aih]	[aj]	[aj]
*-u	[oʔ]	[aw]	[əw]	[əw]	[əw]	[əw]	[əw]	[auh]	[eaw]	[aw]
*-ay	[e]	[a:j]	[aj]	[aj]	[aj]	[aj]	[aj]	[aih]	[aj]	[a:j]
*-aw	[o]	[a:w]	[aw]	[aw]	[aw]	[aw]	[aw]	[aw]	[iəw]	[a:w]

Languages:
(1) = Kayan (Uma Juman),
(2) = Melanau (Serikei)
(3) = Melanau (Dalat, Kampung Kekan)
(4) = Melanau (Dalat, Kampung Teh)
(5) = Melanau (Matu)
(6) = Melanau (Mukah)
(7) = Melanau (Balingian)
(8) = Miri
(9) = Narum
(10) = Kiput[b]

a. The great majority of languages in Sarawak that show vowel breaking are coastal or near-coastal. However, Sa'ban, in the far interior, shows a similar phenomenon that is much less constrained by specific conditions (Blust 2001a). Because of its complexity and separate treatment elsewhere it will not be considered further here.

b. Some irregularities occur. In Serikei Melanau, for example, reflexes of *-uk were recorded as both [uəʔ] and [uək], presumably due to a change in progress. Similarly, although all of the reflexes indicated in Table 8 are well-attested in my data, some forms do not show the expected change. Allomorphs conditioned by low vowel fronting (Blust 2000) are ignored for purposes of this table.

Perhaps the best way to deal with limited data from so many languages is to concentrate our attention on one of these and to extract everything that can be extracted from the available material. Since a published sketch exists for Mukah Melanau (Blust 1988), it can represent the wider collection of languages in Table 8, keeping in mind, however, that each language differs from the others in certain details.

A change that Mukah shares with many of the other languages in Table 8 is for high vowels to undergo breaking with a mid-central offglide before word-final *k and *ŋ. A notable feature of this development is that breaking never occurs before word-final *g, a detail that was checked carefully in Mukah, Dalat (Kampung Kekan, Kampung Teh), Matu, and Balingian, all of which permit word-final voiced stops, unlike Uma Juman Kayan, Miri, Narum, and Kiput. Serikei Melanau has voiced stop finals, but instances of *-g* were not recorded. A second notable feature of this development is that in four of the languages cited here, *i undergoes breaking before word-final *k and *ŋ, but *u either undergoes lowering (1, 3, 5), or shows no change at all (7). Third, Mukah Melanau (MM) shows raising and breaking of *a before final *k, *ŋ, and *R (> *h*). Dalat Melanau, Kampung Kekan (KK) shows raising without breaking, but only before *-k and *-ŋ (which then disappeared). This is seen in (11a), which illustrates these developments, and (11b), which illustrates the absence of breaking before word-final /g/ for the dialects of Mukah and Dalat Melanau and Kampung Teh (KT), where few etymologies are available:

(11) a. *anak > MM [áneəʔ] 'child', *əsak 'cooked, ripe' > [seəʔ] 'red' (cf. [íseəʔ] 'cooked'), *tulak > [túleəʔ] 'push, shove', *bataŋ > [báteəŋ] 'trunk, log', *qudaŋ > [údeəŋ] 'shrimp, lobster', *sandaR > [sádeəh] 'lean against'
*anak > KK [aníʔ] 'child', *m-əsak 'cooked, ripe' > [misíʔ] 'to cook', *tulak > [tulíʔ] 'to push, shove', *hasaŋ > [así] 'gills', *bintaŋ > [bití] 'star', *tuqəlaŋ (> tulaŋ) > [tulí] 'bone'

b. MM [dúhig] 'mythological forest monster', [hig] 'budge, move slightly', [pádʒug] 'foot', [tug] 'ball of the heel', [úug] 'rub, scrub'
KT [lilíg] 'tree resin', [ribíg] 'pinching', [uúg] 'rub, scrub', [tug] 'heel of the foot', *tugtug (> *tutug) > [tutúg] 'to pound (food substances)'

As can be seen from Table 8, final *k has become Mukah glottal stop since breaking took place. By contrast, high vowels lowered before an original final glottal stop, as in *piliʔ > [píleʔ] 'choose', *tumbuʔ > [túboʔ] 'to grow, sprout'.

Two things are particularly surprising about this change. First, it occurred in an unstressed syllable, since Mukah stress is generally penultimate (Blust 1988: 178–179), and second, it happened before word-final [k], [ŋ] and [h] (from *R, which probably was uvular), but not before [g], where high vowels invariably maintain their canonical values of [i] and [u].[7] Other languages show patterns of breaking that are narrower or wider, the narrower patterns favoring breaking

7. Data were gathered in a survey of over 40 languages communities in central and northern Sarawak during an 8-month period in 1971. In some coastal languages, stress was final in citation forms but penultimate in phrasal context.

of *i but not *u, and the wider patterns including many more environments for breaking, as in Serikei Melanau.

The last feature that merits some comment is the mirror-image patterning of breaking in open and closed syllables. In Mukah and a number of the other languages represented in Table 8, high vowels have developed a mid-central *onglide* if they were word-final, but a mid-central *offglide* if they were followed by a triggering consonant. None of this yields transparent cues to a possible phonetic motivation, and the fact that vowel breaking is an areal phenomenon in coastal Sarawak that cuts across languages in different subgroups further suggests that regardless of how it began, at least certain features of this innovation are products of linguistic convention in a context of borrowing rather than products of unconscious, phonetically-driven change.

3.5 Case study 5: Lowering of *u before high vowels in Sula

The Sula language is spoken in two major dialects: Mangon, on the island of Mangole; and Sanana, on the smaller island of Sanana in the Sula Archipelago of eastern Indonesia. The following data have been made available courtesy of Tobias Bloyd, who is currently writing a dissertation on this language in the Department of Linguistics at the University of Hawai'i and has presented many of the following examples in Bloyd (n.d.).

One of the most intriguing features of Sula historical phonology that Bloyd has found through reference to Bahasa Indonesia (his contact language with the community) and comparative data from other languages (Blust & Trussel 2010-) is that *u (but not *i) has lowered to /o/ in the penult if the final syllable contains a high vowel (either *i or *u). Although it will not be pursued further here, after this innovation in a language ancestral to the Mangon and Sanana dialects, *-u then changed to *-a* in Sanana if the penult contained *o, giving rise to such differences as *kutu > Mangon *kotu*, Sanana *kota* 'hair louse' and *susu > Mangon *sosu*, Sanana *sosa* 'breast'.

Table 9 contains data illustrating this change and control cases showing how it is restricted by clear phonological conditions.

Table 9. Lowering of *u before high vowels in Sula[a,b]

1. Lowering of *u		
	PMP	**Sula**
	*tunu	*donu* 'to burn'
	*tuktuk (> *tutuk)	*dotu* 'to strike'
	*bubu	*fofu* 'cage trap for fish and eels'
	*buku	*foku* 'node in bamboo, sugarcane'
	*bulu	*foo* 'hair, feathers'
	*kutu	*kotu* 'hair louse'
	*puluq	*poo* 'ten'
	*pusuq	*poo* 'banana inflorescence'
	*susu	*sosu* 'female breast'
	*ma-putiq	*boti* 'white'
	*kulit	*koli* 'skin'
	*qutin	*oti* 'penis'
	*puki	*poki* 'vagina'
2. Control cases		
	PMP	**Sula**
	*buaq	*fua* (S) 'fruit'
	*tuqa	*ma-tua* 'old (people)'
	*qudaŋ	*uha* (S), *ua* (M) 'shrimp, lobster'
	*Rumaq	*uma* 'house'
	*quzan	*uya* 'rain'
	*pitu	*ga-pitu* 'seven'
	*ŋisi 'bare the teeth'	*nihi* (S) 'tooth'
	*isi	*n-isi* (M), *ihi* (S) 'flesh; contents'
	*niuR	*nui* (< met.) 'coconut'
	*inum	*winu* (M), *win* (S) 'to drink'

a. Unmarked data and data marked (M) come from the Mangon dialect. Sanana forms are marked (S), and are listed where no reflex was found in Mangon.

b. The change *t > *d* in the verbs 'to burn' (*donu*) and 'to strike' (*dotu*) evidently was conditioned by a nasal which was either a proclitic subject marker, or part of an earlier verb prefix (hence pre-Sula *N-tunu, *N-tutuk). The vowel sequence *-oo* in *foo* 'hair, feathers', *poo* 'ten' and *poo* 'banana inflorescence' may be *-ou*, since Sanana has cognates *foa*, *poa*, *poa*.

Both assimilation and dissimilation are necessarily-conditioned sound changes, but their nomenclature implies a reverse symmetry that may be misleading. Assimilations seem clearly to be products of effort-reduction, as in nasal place

assimilation, palatalization, or intervocalic voicing, but nothing similar is discernible in dissimilatory changes. While avoidance of multiple marked structures within the same phonological domain can account for some widespread patterns of dissimilation and thus amount to a type of effort reduction (Blust 2012), no such avoidance can explain many other changes, including the cross-linguistically frequent pattern of liquid dissimilation.

Two proposals intended to explain the pervasiveness of dissimilation in the world's languages are: (1) assimilation facilitates production, while dissimilation facilitates perception (Fromkin & Rodman 1993: 246-); and (2) assimilation is a primary phonetic process, while dissimilation is a reactive reinterpretation in which the speaker *assumes* an assimilation that did not take place and then "undoes" it by subtracting the incorrectly inferred assimilatory contribution (Ohala 1981). Both theories work well for certain types of dissimilation, but do little to help with others, as Eastern Polynesian labial dissimilation, where *fafV > *wahV (but *fofV and *fufV > *hohV and *huhV; cases of *fifV are unknown. Neither theory offers a straightforward explanation for the Sula data. Because lowering was triggered by both *i and *u but affected only *u, this change cannot be characterized exclusively as the dissimilation of identical segments. This is unusual in itself, since nearly all reported dissimilations eliminate segmental repetition. To make matters still more theoretically recalcitrant, since *i was not affected, the lowering of *u before high vowels in Sula cannot even be characterized uncontroversially as a more abstract process of height dissimilation.

How this change might facilitate the listener's perception of the speech signal, or alternatively how it might have arisen by "undoing" a perceived assimilation, is difficult to see, since there is no known evidence that *oCu* or *oCi* is easier to perceive than *uCu* or *uCi*, nor is there any evidence that an assimilation process ever existed in the language raising /o/ to /u/ where the next syllable contained a high vowel.[8]

8. One referee notes that my account of the apparent dissimilation of *u before high vowels in Sula "does not take into account more recent work on asymmetries. Cole et al. (2011) contains a good survey, and Beddor et al. (2002) find exactly the asymmetry between i and u that is assumed to be problematic in the discussion here of Sula." Each of these studies is concerned with vowel-to-vowel coarticulation, while the Sula data involve an apparent height dissimilation in which H-H became M-H for *u but not *i. For vowels of identical height co-articulatory effects are expected for F2 (frontness), but for not F1 (height), so without a more developed argument it is hard so see how this suggestion can contribute anything to our understanding of the peculiar change in Sula.

3.6 Case study 6: Low vowel dissimilation in Oceanic languages

Another vowel change that appeals to unexpected conditions is also dissimilatory, but involves raising rather than lowering of a penultimate vowel. This occurs in Oceanic, a well-defined subgroup of about 460 Austronesian languages in the Pacific regions of Melanesia, Micronesia and Polynesia.

Bender (1968) noted that in Marshallese of eastern Micronesia, *aCa dissimilated to *eCa* after loss of the final vowel had reduced unsuffixed bases to monosyllables, and this has given rise to synchronic alternations in various independent nouns and their "construct" forms indicating '[noun] of', as with *bar*: *bera-n* 'head', *yat*: *yeta-n* 'name', or *maj*: *meja-n* 'eye'. Similar examples of low vowel dissimilation (LVD) have been reported for other Oceanic languages in the Admiralty Islands of western Melanesia (Blust 1996) and in Vanuatu (Lynch 2003), and also in three genetically unrelated languages in other parts of the world, including Alamblak, a Sepik Hills language of New Guinea, Kera, an East Chadic language in Chad and Cameroon, and southern Russian (Blevins 2009, citing Odden 2005 for the latter two cases). Table 10 provides evidence of this change in Ere, an Oceanic language spoken near the eastern end of Manus in the Admiralty Islands of Papua New Guinea.

Table 10. Low vowel dissimilation in Ere of Eastern Manus Island

POC[a]	Ere
*mata-ña	*mira-n* 'his/her eye'
*na tamata	*ndimat* 'person, human being'
*kamali	*kimal* 'men's house; male'
*pataŋ	*pira-n* 'its stem'
*sa-ŋa-Ratus (> *saŋat)	*siŋat* '100'
*katapa	*kirah* 'frigate bird'
*kanase	*kinas* 'mullet'
*lapat	*lipa-n* 'big, great; God'
*ŋajan	*ŋira-n* 'his/her name'
*tama-ña	*tima-n* 'his/her father'
*salan-ña	*sila-n* 'its path'
*panako-akin	*pinaw-ey* 'to steal (trans.)'

a. POC = Proto-Oceanic

As in Marshallese, final vowels appear to have been lost prior to the appearance of low vowel dissimilation, giving rise to alternations such as *sal* 'path, road': *sila-n* 'its path'.

At least two things are puzzling about this change. First, since Blevins (2009) has found parallel cases in three other genetically disparate languages it is difficult to escape the conclusion that LVD is phonetically motivated. This conclusion is made even more compelling by the observation that in both Oceanic languages and in the East Chadic language Kera, LVD is blocked by an adjacent laryngeal consonant. However, it remains unclear what this motivation might be, since, as noted earlier, *a* is the one vowel that all languages must have (the so-called "prince of vowels"), an observation that makes it difficult to see why it would need to be altered if it was followed in the next syllable by a copy of itself.

It is perhaps equally puzzling that this change is found in widely-separated Oceanic languages (the Micronesian, Manus and Vanuatu cases must be considered historically unconnected), yet it occurs nowhere else in the roughly 800 Austronesian languages outside the Oceanic group.

4. Odd conditioning of consonant changes

The following six case studies focus on consonant changes, some of which are surprising in themselves, but that are made even more surprising by the conditions under which they occur. The first of these is the backing of labials to velars in the Atayalic languages of northern Taiwan. While changes such as *-n > -*ŋ* are crosslinguistically common, the Atayalic languages backed both stops and nasals to velars word-finally if they were labial (but not if they were dental). The second is the change of intervocalic *b to *-k-* or *-kk-* in the Berawan languages of northern Sarawak (Malaysian Borneo), a change that probably occurred in two steps: (1) *-b- > *-g-*; (2) intervocalic devoicing. Both of these changes are contrary to what phonological theory would lead us to expect. The third case study shows the peculiar development of the voiced velar stop *g in Sa'ban of northern Sarawak, which became *j* initially and intervocalically (regardless of the following vowel), but regularly became *p* word-finally and occasionally word-initially. The fourth is the gemination of an open final syllable onset in the Berawan languages of northern Sarawak. The fifth is the gemination of final consonants before a historically secondary supporting vowel in Talaud of northern Sulawesi, Indonesia. This unusual gemination process raises the question whether complex sound changes (changes that affect more than one segment simultaneously through unrelated phonological processes) are real. The last case study of consonant change under theoretically unexpected conditions involves glide fortition in Narum of northern Sarawak. While glide fortition is common cross-linguistically, Austronesian languages differ in whether they allow it only for phonemic glides or for all phonetic glides,

even where these are non-contrastive. Narum is odd in allowing glide fortition only for non-contrastive glides, while leaving phonemic glides unchanged.

4.1 Case study 7: Backing of final labials to velars in Atayalic

In many languages, word-final consonants are subject to various processes of lenition that may ultimately result in loss. However, stops and nasals normally follow very different trajectories: voiced stops tend to devoice, and the entire series of word-final voiceless stops then merges as glottal stop, while *-m and *-n tend to merge as *-n*, and the alveolar nasal from both sources then falls together with its velar counterpart. Li (2004a, b), however, has shown that the Palŋawan dialect of Atayal and all dialects of the closely related Seediq of northern Taiwan change labials to velars word-finally, whether they are nasals or stops. More exactly, final *b and *p first merged as *-p, which then shifted to a velar, as shown in (12), where PA = Proto Atayal and PS = Proto-Seediq, each of which represents a coordinate branch of the Atayalic group of Austronesian languages, descending from Proto-Atayalic (PAT):

(12) PA *h<um>agub, *hagab-an > Palŋawan *h<um>uk*, *hub-an* 'to do magic'; PA *ma-gagub, *pa-gagub-an > Palŋawan *ma-gaguk*, *pa-gagub-an* 'to share one cup', PAT *ma-surab > PS *mə-surak 'to yawn'
PAT *m-iyup > PS *m-iyuk 'to blow'; PAT *gəhap > PS *gəhak 'seed'

Non-final labials did not undergo this change, as can be seen by the alternations in the suffixed forms of Palŋawan, or by etymologies such as PAT *babaw > PS *babaw 'above', PAT *tubaʔ > PS *tubaʔ 'fish poison', PAT *hapu-niq > PS *pu-niq 'fire', PAT *ma-pituʔ > PS *mə-pituʔ 'seven', PAT *gipun > PS *gupun 'tooth', or PAT *s<um>apag > PS *s<əm>apag 'unroll a mat'. Since *-m > *-ŋ* is attested in the same set of languages, it is clear that the backing of labials to velars word-finally was a general process, but what property of word-final position would condition this change remains obscure. As already noted, parallels exist for the backing of word-final nasals to *-ŋ*, as in Mandarin Chinese or languages of southern Sulawesi such as Buginese and Makasarese, but neither a phonetic motivation nor any parallel in other languages is known for the backing of labial stops to velars in the same position.

4.2 Case study 8: Intervocalic labial to velar stop in Berawan

Blust (2005) shows that some of the languages of northern Sarawak have devoiced intervocalic obstruents. While this is unexpected in itself, it is only part of what is perhaps the most striking form of this change, which converted *b to *-k-* or *-kk-* in the Berawan dialects, as shown in Table 11 for the dialect of Batu Belah

(PNS = Proto-North Sarawak). The second odd feature of this change, gemination of the onsets of open final syllables, will be treated below. For a detailed account of Berawan historical phonology see Burkhardt (2014, 2015).

Table 11. Shift of *b to -k- in Batu Belah Berawan

PNS	Batu Belah
*abaŋ	*akiŋ* 'ilipe nut'
*abu	*akkuh* 'ashes'
*balabaw	*bəlikiw* (LT: *bəlio*) 'rat'
*babuy	*bikuy* (LT: *biwih*) 'pig; boar'
*bubu	*bukkuh* 'fish trap'
*bubuŋ	*bukuŋ* 'ridge of roof'
*Ribu	*gikkuh* 'thousand'
*Rabun	*gikuŋ* (LT *gium*) 'cloud, fog'
*kabiŋ	*kakiŋ* 'left side'
*lubaŋ	*lukiŋ* 'hole'
*mabuk	*makuk* (LT: maoʔ) 'drunk'
*nibuŋ	*nikuŋ* (LT: niuŋ) 'nibong palm'
*tuba	*tukkih* 'derris root'
*ubi	*ukkih* 'yam, cassava'

It seems reasonably clear that *b > *-k-* in the Berawan dialects involved two ordered steps: (1) *b > *-g-* and (2) intervocalic devoicing. Although both changes are inextricably intertwined in this reflex, they can be separated in the parallel change of *R (probably a uvular rhotic) to initial *g* but medial *k,* since here the place features agree but the voicing feature does not, as shown in (13):

(13) PNS *Ratus > BB *gitoʔ* '100', *Ratas > *gitaʔ* 'milk', *Rusuk > *gusok* 'chest', but PNS *aRəm > BB *akəm* 'pangolin, scaly anteater', *baRa > *bikkeh* 'shoulder', *duRi > *dukkih* 'thorn', *mə-baRaʔ > *kiki* 'swollen', etc.

Since *p was unaffected by this change (PNS *apuy > BB *apoy* 'fire', PNS *səpaʔ 'to chew (betel nut)' > BB *səpa* 'to chew', PNS *kapal > BB *kapan* 'thick, as a board', PNS *tapan > BB *tapan* 'winnowing basket', etc.), intervocalic devoicing must have followed the backing of the medial labial stop. The first step in this unusual change, then, was the backing of *b (but not *p) only in intervocalic position. Again, there is no obvious connection between the phonetic properties of the segment undergoing change and those of the conditioning environment, any more than there is between word-final position and the backing of *-p in Palŋawan Atayal or Seediq as described in the previous section.

4.3 Case study 9: *g > *p-, -j-, -p* in Sa'ban

Sa'ban is a highly innovative dialect of Kelabit in the phonological "hot spot" of northern Sarawak, where several languages show extensive and sometimes quite odd sound changes (Blust 2001a). Among many strange innovations in this language is the conversion of *g to *j* ([ʤ]) intervocalically and initially in some forms, but to *p* word-finally and word-initially in two forms, as seen in Table 12.

Table 12. *g to *j-, -j-* and *g to *p-, -p* in Sa'ban of northern Sarawak

1. *g to *j-, -j-*		
	PKLD[a]	**Sa'ban**
	*gatəl	*jatəl* 'itchy'
	*gənuluh	*jənləw* 'empty rice head'
	*gəramih	*jəlaməy* 'rice straw'
	*gituʔən	*jɪntoʔɔn* 'star'
	*agag	*ajiəp* 'rice sieve'
	*ŋ-agap	*m-ajeəp* 'to startle'
	*sagət	*ajɪt* 'quickly'
	*iguʔ	*jəuʔ* 'shame'
	*pəgamuŋ	*pəjamuəŋ* 'tangled'
	*təgəkər	*təjaʔəl* 'to shiver'
	*təgəraŋ	*təjareəŋ* 'ribs'
2. *g to *p-, -p*		
	PKLD	**Sa'ban**
	*gərawat	*pəlawət* 'complicated'
	*giləg	*peləp* 'skittish'
	*agag	*ajiəp* 'rice sieve'
	*ŋ-agap	*m-ajeəp* 'to startle'
	*arəg	*arəp* 'crumbs, rubbish'
	*bəluʔug	*bəlʔup* 'wasp, hornet'
	*iləg	*eləp* 'to separate, divorce'
	*ələg	*ləp* 'to stop, as working'
	*pəpag	*ppap* 'a slap'

a. (PKLD = Proto-Kelabit-Lun Dayeh).

In a number of other forms *g was lost word-initially, as with *gaiŋ > *ayəŋ* 'spinning top', *garəŋ > *arəŋ-arəŋ* 'to roar', *gətiməl > *hməl* 'bedbug', *guliŋ > *liəŋ liəŋ* 'shake head sideways', or *guta > *toə* 'cross a river'; occasionally other, apparently

unique developments appear, as with *giwən > *lawən* 'messy, untidy', or *m-alug > *malok* 'to trick'. It is important to note that Sa'ban is closely related to Kelabit, which shows none of these changes. Since about 83% of the vocabulary on a Swadesh 200-word list is cognate between Sa'ban and the Bario dialect of Kelabit, the local standard, one could even call these dialects of the same language, despite the reaction of Kelabit speakers that Sa'ban is a language that is "totally different" from theirs. The changes that affected this language community must, therefore, have been both drastic and rapid. Since PKLD *b and *g merged as Sa'ban *-p,* it is possible that *g > *-p* was a two-step change *g > *-b > *-p*. However, the two known cases of *g > *p* word-initially suggest that this was a one-step change, since *b never devoices word-initially. Notably, although *g underwent what appears to have been an unconditioned phonemic split in initial position, it is never reflected as *p* intervocalically or as *j* word-finally.

4.4 Case study 10: Gemination of an open final syllable onset in Berawan

The Berawan languages of northern Sarawak are notable for a number of exotic sound changes, suggesting that phonetically unmotivated change may be more common in some languages than others for cultural reasons that remain to be elucidated. Another change common to all four dialects of Berawan (Long Terawan, Long Teru, Batu Belah, Long Jegan) is the gemination of the onset of an open final syllable (later closed by addition of *-h*), as shown for Long Terawan in Table 13.

Table 13. Gemination of the onset of an open final syllable in Long Terawan Berawan[a]

PNS	LTB
*lima	dimməh 'five'
*Ribu	gikkuh '1000'
*ulu	ulloh 'head'
*bulu	bulluh 'body hair, feather'
*tawa	tabbəh 'to laugh'
*mata	mattəh 'eye'
*laki	lakkeh 'male, man'
*bana	binnəh 'husband'
*tina	tinnəh 'mother'
*tama	tamməh 'father'
*tali	talleh 'rope'
*lipa	lippəh 'snake'
*batu	bittoh 'stone'

Table 13. (*continued*)

PNS	LTB
*abu	akkuh 'ash'
*kayu	kajjuh 'wood, tree'
*ubi	ukkih 'yam'
*duRi	dukkih 'thorn'

a. Since most consonants automatically geminate after penultimate schwa, etymologies such as *təlu > *təlloh* 'three' are omitted here.

The onset of closed final syllables did not geminate (PNS *lakaw > LTB *lakaw* 'to walk, go', *apuk > *apoʔ* 'dust', *tulaŋ > *tulaŋ* 'bone', *anak > *anaʔ* 'child', *utin > *utin* 'penis', etc.), implying that the addition of *-h* postdated gemination but predated the loss of word-final reflexes of PNS *ʔ and *R, since these blocked gemination before they disappeared, as shown in (14):

(14) PNS *tujuʔ > LTB *tusu* 'seven', *puluʔ > *pulo* '10', *jəlaʔ > *jəli* 'tongue', *tanaʔ > LTB *tana* 'earth, soil', *lasuʔ > *laso* 'warm, hot', *putiʔ > *pute* 'white', *mataʔ > *mata* 'raw', *dəd^{h}uR > *dicu* 'woman', *ikuR > *iko* 'tail', *tuduR > *tudo* 'to sleep', etc.

Again, a phonetic connection between change and environment is elusive. Gemination in Austronesian languages has generally arisen in one of four ways: (1) by assimilation of a nasal to a following homorganic stop, (2) assimilation of heterorganic consonants in a cluster, (3) compensatory lengthening after schwa followed by merger of schwa with some other vowel, or (4) syncope between identical consonants or consonants that share a common place feature (Blust 2013: 648–651). If conditioned by syllable type, one might expect the onset of a *heavy* syllable to geminate, as in the Papuan language of West Pantar (Heston n.d.), not the onset of a light syllable. Word-final vowels are lengthened automatically in many of the languages of Sarawak, including Berawan, but this was recorded only in citation forms, and there is – at least initially – no obvious reason why lengthening a word-final vowel would favor the lengthening of its syllable onset.

An anonymous referee has pointed out that Gordon (1997) may have provided a solution to this problem. Most Balto-Finnic and Saamic languages exhibit a system of "consonant gradation" in which a more fortis member of the inventory appears in some environments, alternating with a more lenis member in others (*pappi* vs. *papista*, *lapa* vs. *lavasta*, etc.). Details differ across languages, but the basic claim is that foot-final vowels in unstressed open syllables were automatically lengthened, and that vowel lengthening triggered a concomitant lengthening of the preceding consonant to offset the additional weight of the unstressed syllable. Gordon argues that this process of onset fortition results from a universal

tendency in languages to avoid prosodic feet in which the stressed syllable is lighter than the unstressed syllable.

This analysis departs from the traditional view in Finnic linguistics, in which consonant gradation is seen as a product of lenition in closed syllables rather than fortition in open ones. Gordon acknowledges that his interpretation complicates the explanation for the situation in modern Finnish, but simplifies the description of the relevant facts in Sámi, which he takes – contrary to established opinion – as reflecting a more conservative state of affairs than Finnish. Most importantly in Gordon's view, the new analysis is better supported by considerations of general phonological plausibility.

It remains to be seen how this proposal will be received by the majority of specialists in Finnic linguistics, but for now we need to know how well it fits the facts in Berawan. Like Balto-Finnic languages, most Austronesian languages that do not have phonemic stress have trochaic feet. However, this does not appear to be the case in Berawan. Wherever I recorded stress in my fieldnotes, both in citation forms and in sentence context for the dialects of Long Terawan and Long Jegan, it was consistently word-final. This agrees with the independent conclusions of Clayre (1996) and García-Bellido & Clayre (1997), the first of whom noted that "The minimal word in Berawan consists of one nuclear syllable (S)" (i.e. a monosyllable). She goes on:

> The maximum expansion of a word consists of a nuclear syllable preceded by one or two pre-nuclear syllables (s) … Stress in Berawan falls on the nuclear syllable (S). It is marked by length and by higher, often rising, pitch, and not by high average amplitude. The same pattern was also recognized in sentence utterances (see García-Bellido and Clayre 1997, § 4.1.3. and Figure 5b). Clayre (1996: 212)

It seems reasonable to conclude, then, that prosodic feet in Berawan are not trochaic, but iambic. In Balto-Finnic languages the foot is left-aligned, and if this were also true for Berawan it might be easier to apply Gordon's solution to this data, since a final open syllable would remain unstressed and long, even if it was extrametrical rather than foot-final. But feet in Austronesian languages generally, whether they are trochaic or iambic, are right-aligned, meaning that an open final syllable in Berawan is necessarily stressed. To salvage Gordon's appeal to a proposed language universal, we might assume that Berawan languages had trochaic feet at the time the onset of open final syllables was geminated. However, Burkhardt (2013) reconstructs Proto-Berawan stress on the ultima based on a systematic comparison of all four modern dialects (Long Terawan, Long Teru, Batu Belah, Long Jegan).

Perhaps more seriously, even if we assume that Proto-Berawan had trochees where we now find iambs because of later areal developments in Sarawak, in

applying Gordon's proposed language universal to Berawan we must ask why the same gemination did not happen in scores of other Austronesian languages that are known to have unstressed foot-final open syllables with allophonically long vowels or in other language families. It goes without saying that language universals need not be expressed in every language, but even with statistical universals, one expects a distribution that departs from chance. As noted in Blust (1980: 286) with regard to optimal explanations for the distribution of non-linguistic culture traits, a universalist explanation is plausible only where the trait in question has a wide and scattered distribution among linguistically unrelated groups, the classic example in language perhaps being the distribution of *mama* and *papa/baba* as parent terms. In the case at hand, this condition does not hold, even if we assume against the current evidence that the onset of open final syllables was geminated because its nucleus was unstressed at the time of this peculiar change. True language universals are simply more richly attested than this, and given the number of languages and their dialects in which phonological processes can be found, it is hard to rule out the possibility that the apparent parallelism of fortition processes in Balto-Finnic languages and Berawan is anything more than a product of chance convergence.

This analysis is initially appealing, then, but on closer inspection we are left with the question we faced at the beginning: why would a language geminate the onset of open final syllables and no others?

4.5 Case study 11: Gemination of final consonants before a supporting vowel in Talaud

The Berawan languages of northern Sarawak show gemination in an unusual environment, namely the onset of open final syllables. Talaud, spoken in the Talaud Islands between Sulawesi in central Indonesia and the southern Philippines, shows gemination in another unusual environment, since here it affects original final stops that have been placed in intervocalic position through the addition of a paragogic vowel, as seen in Table 14 (Sneddon 1984, Blust 2012).

Like some other innovations examined earlier, this change involves two inseparable elements: (1) consonant gemination and (2) a paragogic vowel. The question thus arises whether this was a single change or a sequence of two changes. Sneddon (1984: 48) notes that for both phonetic and phonotactic reasons, "It is unlikely that gemination occurred word-finally so it may be assumed that it developed after addition of the paragoge". However, as seen in (15), words that ended in *-Ca did not geminate the onset of the final syllable:

Table 14. Gemination of original final consonants in Talaud

PMP		Talaud	PMP		Talaud
*-p	>	*-ppa*	*qatəp	>	*atuppa* 'roof, thatch'
*-t	>	*-tta*	*Ramut	>	*žamutta* 'root'
*-k	>	*-ʔa*	*anak	>	*anaʔa* 'child'
*-b	>	*-bba*	*tutub	>	*tutubba* 'to close, as a door'
*-d	>	*-dda*	*likud	>	*liʔudda* 'back'
*-g	>	?			
*-m	>	*-mma*	*inum	>	*inumma* 'drink'
*-n	>	*-nna*	*taqun	>	*tonna* 'year'
*-ŋ	>	*-ŋŋa*	*dadaŋ	>	*daraŋŋa* 'warm by a fire'
*-s	>	*-ssa*	*Ratus	>	*žasutta* '100'[a]
*-l	>	*-lla*	*bəŋəl	>	*beŋella* 'deaf'
*-R	>	*-kka*	*niuR	>	*niukka* 'coconut'

a. Like several other languages in both the northern Philippines and northern Sulawesi, Talaud shows regular metathesis of *tVs or *tVCVs (Blust 2013: 642–43).

(15) *apa > *apa* 'what?'
*lima > *lima* 'five'
*mata > *mata* 'eye'
*qasawa > *sawa* > *saβa* 'spouse'
*tiŋada > *tiŋara* 'look upward'

It appears then that final geminates either existed before paragoge in etymologies like *qatəp > *atuppa* 'roof, thatch', or that gemination and paragoge were a single change. But what would motivate consonants to geminate word-finally, where they would be maximally difficult to distinguish from their singleton counterparts? And if this solution is rejected, what would motivate consonant gemination only as part of a complex sound change in which *-a* was added as a paragogic vowel? Again, the relationship between gemination and its conditioning environment is just as obscure here as it is for gemination in a very different odd environment in the Berawan languages of northern Sarawak.

4.6 Case study 12: Atypical glide fortition in Narum

Glide fortition is a common type of sound change in Austronesian languages, being widely attested in the languages of Borneo and over a very large geographical area elsewhere, including the Aru islands of eastern Indonesia, Chamorro of

western Micronesia, and various parts of Melanesia, reaching from the Admiralty Islands in the Bismarck Archipelago to north-central Vanuatu (Blust 2013: 612–614). Languages that have undergone this process fall into two types: those that strengthen only phonemic glides, leaving allophonic glides untouched, and those that strengthen phonetic glides, whether they are phonemic or not, as shown with Lau (type 1) and Chamorro (type 2) in Table 15.

Table 15. Two patterns of glide fortition in Austronesian languages

PMP	Lau	Chamorro	gloss
*wada	–	*gwaha*	'have; there is'
*wahiR	*kwai*	–	'fresh water'
*walu	*kwalu*	*gwalu*	'eight'
*qawa	*ʔakwa* 'fish sp.'	*agwa*	'milkfish'
*lawa	*lakwa*	–	'spider'
*siwa	*sikwa*	*sigwa*	'nine'
*duha (> *dua*)	*rua*	*hugwa*	'two'
*lahud (> *laud*)	*lau*	*lagu*	'seaward; North'
*buaq	*fua-*	*pugwaʔ*	'fruit; betel nut'
*zauq	*tau*	*chagoʔ*	'far, distant'
*layaR	–	*ladzak*	'sail'
*niuR	*niu*	*nidzuk*	'coconut'
*tian	*īa*	*tudzan*	'belly'

Although Lau also strengthened *y to *s*, a historically-related type of front glide fortition is shared with other languages of the southeast Solomons that do not strengthen *w, and it therefore appears to be independent of the issue addressed here (Lichtenberk 1994). Regardless of the status of *y > *s*, it is clear from the above material that both Lau and Chamorro strengthened *w to a labiovelar stop before a non-rounded vowel (*kw* in Lau, *gw* in Chamorro), and that Chamorro strengthened *w to *g* before *u/o*. But the two languages differ in the form of glide fortition with regard to the automatic transitional glides [w] (in the sequences [uwa] and [awu]), and [j] (in the sequences [iju] and [ija]): Chamorro strengthens these predictable glides just like *w and *y, while Lau does not. There are various ways to look at this divergent evolution, the most obvious one being that it happened on the phonetic level in Chamorro, but the phonemic level in Lau, where it would thus presumably have been more conscious. Whatever analysis one favors in explaining this difference, which can be multiplied with many other Austronesian

cases, most of which pattern like Chamorro, nothing from the history of other languages can prepare us for the patterning of glide fortition in Narum.

Narum, spoken around the market town of Marudi, about 50 miles inland on the Baram River of northern Sarawak, is another member of the phonologically highly innovative North Sarawak group of languages. Like Chamorro and many other Austronesian languages, it strengthens the automatic transitional glides [w] and [j] in phonemic sequences such as /ua/ or /ia/, so that [w] has become /b/ and the palatal glide [j] has become /j/ (a voiced palatal affricate). However, what is totally unexpected in view of this development is that it leaves phonemic glides unchanged, as shown in Table 16.

Table 16. Divergent development of non-phonemic and phonemic glides in Narum of northern Sarawak

1. Non-phonemic glides	**PMP**		**Narum**
	*baRuaŋ ([baɣuwaŋ])		*bərəbiəŋ* 'Malayan sun bear'
	*quay (> *uay*; [uwaj])		*bi* 'rattan'
	*buat ([buwat])		*biət* (< *wat) 'long'
	*duha (> *dua*; [duwa])		*dəbeh* 'two'
	*Ruab ([ɣuwab])		*rəbiəp* 'high tide; high (of the river)'
	*laqia (> *lia*; [lija])		*ləjeəh* 'ginger'
	*tian ([tijan])		*tijiən* 'belly, abdomen'
2. Phonemic glides	**PMP**	**PNS**	**Narum**
	*qasawa		*awəh* 'spouse; wife'
	*lawaq		*ha-lawaʔ* 'spider'
		*jaway	*jaweəy* 'face'
		*pawat	*pawat* 'large fruit bat'
	*buqaya (> *buaya* > *baya*)		*bayeəh* 'crocodile'
	*daya		*dayeəh* 'interior, away from river'
	*kahiw (> *kayu*)		hayeəw 'wood, tree'
	*Raya		(r)ayaʔ 'big'
		*kuyad	*uyit* 'gray langur'

Malay loans, which presumably entered the language after this change was complete, show no change to phonemic glides and do not develop obstruents medially from predictable glides: *bawiəŋ* < Malay *bawaŋ* 'onion', *bayiər* < Malay *bayar*

'to pay', *layar* < Malay *layar* 'sail', *payuəŋ* < Malay *payuŋ* 'umbrella', *dian* < Malay *dian* 'candle'. All words in Table 16 appear to be directly inherited, and the divergent pattern of glide fortition is very transparent: phonemic glides remained unchanged, but non-phonemic glides were strengthened.

This pattern is particularly difficult to reconcile with theories that require sound change to be either phonetically or phonologically motivated. In Lau, one can argue that the change happened on the level of the phoneme, and so was relatively conscious, unlike the case in Chamorro and most other Austronesian languages, where it can be argued that the change happened on the phonetic level, affecting any glide that was present in speech.

However, Narum presents a more difficult problem. Here we cannot claim that glide fortition is *either* phonetically or phonologically motivated, since if it were phonetically motivated we would expect it to affect all glides in the speech stream, while if it were phonologically motivated we would expect it to affect just phonemic glides, as in Lau. Because neither of these choices is available it is simplest to assume that glide fortition in Narum is a product of areal influence which involved the conscious manipulation of phonetic material without regard to either phonetic or phonological conditioning.[9]

5. Conclusions

As noted at the beginning, in one sense there is nothing new about this article, since linguists have long known that some sound changes, or at least reflexes, involve odd conditions. Indo-Europeanists need only think of the peculiar *ruki*-rule for Sanskrit (or more properly proto-Indo-Iranian), in which *s became a palatal or retroflex sibilant after *i, *u, *r or *k, the change of Proto-Indo-European *dh to /f/ in Latin, or of Latin velars to Romanian labials immediately preceding dentals (*lactem* > *lapte* 'milk', etc.). The first point of this article is to show, with data from a very large language family that is well-studied but unfamiliar to most historical linguists, that a surprising lack of connection between the phonetic properties of segments undergoing change and those of their conditioning environments is not

9. As one referee pointed out, one could, of course, argue that this change is phonetically conditioned in that glide fortition occurred only after high vowels, but not after the low vowel *a. However, since glides are automatically inserted between a high vowel and a following unlike vowel, but not between a low vowel and a following high vowel, these two possible analyses cannot be clearly separated. Regardless of which interpretation one favors, this is a very odd condition for glide fortition, which normally affects all glides whether or not they are contrastive, or only glides in non-predictable environments, hence those that follow low vowels.

as unusual as is commonly assumed; the 12 examples cited here are only the tip of the proverbial iceberg.

It is universally agreed that the most reasonable approach to such puzzling developments is to search for evidence of secondary developments that have affected a primary change that *was* phonetically motivated. In some cases, this has proved to be a productive research strategy, and a clear difference has been shown between the contemporary by-product of change or changes (the reflex), and the primary, phonetically motivated change itself (the actuation); for a particularly clear example of the success of this approach cf. Janda & Joseph (2003: 210–211) on subdialectal differences in the conditions for lowering *o in Schaffhausen German. However, the observation that some surprising reflexes are products of the generalization of the original conditions for a change as it spread from community to community is no guarantee that all surprising reflexes have similarly complex layerings of history.

In some ways, it is hard to decide what to conclude from an examination of the type of data presented here and in Blust (2005), since we simply do not know whether these surprising reflexes are single-step changes lacking phonetic motivation or deceptive products of secondary reworking. I certainly keep an open mind to the possibility that what appears to be the product of a single-step change may have a more complex history, but I cannot accept this type of interpretation as an "explanation" when it is proposed solely to maintain the appearance that all sound changes begin as products of unsurprising phonetic processes.

With reference to the 10 cases presented in Blust (2005), Garrett states that:

> One obvious difficulty with this line of reasoning is that a series of sound changes, each phonetically grounded, can yield an outcome or synchronic pattern that corresponds to no phonetic process; on Blust's dossier see Goddard (2007) and Blevins (2007, 2008b). The error in such cases is to assume that the telescoped result of multiple changes arose through a single sound change; admittedly, it may not be clear when telescoping is plausible rather than a hypothesis of convenience.
>
> Garrett (2015: 232–233)

Given this rush to dismissal one might as well add "Case closed; let's move on to real business." Yet the discussion of sound change in Blust (2005) addresses precisely this issue:

> First, at the risk of belaboring the obvious, the notions "reflex" and "sound change" must be kept apart. A reflex is a mapping of a proto-form onto its historical continuation. This may involve one or more sound changes. Accumulations of sound changes can produce "telescoping" whereby a reflex appears to involve a phonetically unmotivated sound change, but is actually the product of successive natural changes.
>
> Blust (2005: 226)

Each of the case studies in Blust (2005) and in the present article has been examined with this possibility in mind, and for each of them a theoretically orthodox interpretation was rejected as being a product of theory-driven speculation rather than of observation-based analysis. In some cases, my skepticism may have been misguided, but it was intended to "put the brakes" on rapid dismissal of such cases through superficial speculations about possible unseen developments, rather than seriously considering the possibility that in some cases the actuation of change is a product of conscious choice.

A brief examination of the counter-proposals that Garrett mentions in passing, namely Goddard (2007) and Blevins (2007, 2008b [= 2008]), might therefore be the best way to conclude this article, as it returns the discussion to the central issues that I have raised:

1. Are all sound changes or the conditions for them linguistically motivated?
2. When are we justified in substituting speculation for the hypothetico-deductive method in seeking to explain reflexes that violate our assumptions about possible sound changes?

Goddard (2007) specifically addresses glide fortition in western Manus (Admiralty Islands, Papua New Guinea), whereby earlier *w and *y merged as *-p* in Drehet and Levei when they were in historically secondary coda position. The original discussion noted that *y > *-p* is the more recalcitrant change, given the radical featural differences between these segments, and it was suggested that if this reflex is in fact a product of telescoping, a historical sequence of the type *y > *w* > *v* > *f* > *p* presumably would need to be posited. By way of response, Goddard points out (2007: 117) that Proto-Algonquian *w became Proto-Arapaho-Gros Ventre *y unconditionally, and since the merger of labial and palatal glides is attested in one direction (merger as *y*) in one language family, there is a priori no reason to doubt that it could have happened in the opposite direction (merger as *w*) in another. I do not deny the reasonableness of this argument, but the fact remains that in a language family with well over 1,000 members and many more dialects, the merger of *y and *w has never been reported elsewhere, and crucially, in Drehet and Levei, the change was conditioned. Even if we accept a change *y > *w* that is, so far as we know, unique to Austronesian languages, we must still contend with the fact that – unlike the reverse change in Proto-Arapaho-Gros Ventre – this merger happened in some environments but not others, and there is no obvious phonetic relationship between environment and change. There is no need to enter into full details here, but both languages lost Proto-Manus final vowels and underwent fortition of all glides, including those that were earlier automatic transitions between a high vowel and a following unlike vowel, as in *ia [ija] > Drehet, Levei *ip* '3.SG', or *kalia > Drehet, Levei *kalip* 'grouper'. Since *w and *y are reflected without change in

onset position, we are forced to conclude that they merged word-finally after the loss of final vowels before undergoing fortition to *-p*. As pointed out in the original paper, and discussed accurately by Goddard, there may have been structural reasons to introduce a final *-p* into the immediate ancestor of these two closely-related languages: prior to glide fortition, only five consonants were allowed word-finally: *-k, *-ŋ, *-h, *-w and *-y (just *-k, *-ŋ, *-h, and *-w adopting Goddard's hypothesis), and *-w* would have been the phonetically closest source for such a segment. However, there is no obvious reason why *w and *y would merge and then produce *-p when they could have remained distinct and produced both *-p* and *-t* by glide fortition. Under the interpretation favored by Goddard (2007), the original problem, which was posed as one of segmental change, can now be seen as one of "odd conditions": even if we assume a unique merger of *w and *y in a language family with over 1,000 members, why would this merger happen only in coda position, leaving onset glides untouched?

Garrett's dismissal of this particular case because it was provisionally addressed by Goddard (2007) is unfortunately typical of the spirit with which such problematic cases have been treated in the literature: they are theoretically inconvenient and so are best passed over quickly.[10] The fact is, none of the alternative explanations that have been proposed for the data considered in Blust (2001b, 2005) and in this paper stand up to close scrutiny. This includes Gordon (1997) on gemination of the onsets of open final syllables in Berawan; Goddard (2007) on merger of *-w/y as *-p* in Drehet and Levei; Beddor, Harnsberger & Lindemann (2002) or Cole, Linebaugh & McMurray (2010) on *u dissimilation in Sula; or Blevins (2007) on the nasalization of voiced stop codas in Northern Batak and Berawan, although I have since adopted an explanation for this recurrent change which does appeal to phonetic motivation – but not the explanation favored by Blevins. It is discouraging to see survey papers on sound change or comments by referees that hastily dismiss these difficult cases, or claim that they are theoretically tractable by leaving out essential details: gemination in Berawan (wrong foot type, distribution); glide fortition in Drehet/Levei (condition ignored); height dissimilation in

10. Blevins (2007) addresses the nasalization of voiced stop codas in Northern Batak and Berawan, which Blust (2005) treated as phonetically unmotivated. Her suggestion that *-b/d/g were allophonically prenasalized, however, is contrary to everything we have learned about the distribution of prenasalized stops in Austronesian languages over the past century and a half. I have since concluded that voiced stop coda nasalization is simply a less common alternative to final devoicing (Blust 2016). With regard to Blevins (2008), which challenges a claim that stressed vowels were deleted in the history of Mussau (Blust 2001b), she cites Ross (2002) on the position of stress in certain lexical items of modern Mussau because it supports her rejection of this claim, but my transcriptions for these forms (admittedly working with a single speaker) clearly and consistently differ from those of Ross.

Sula (coarticulatory effects irrelevant); and deletion of stressed vowels in Mussau (disagreements in transcription of stress). Collectively, this dismissive approach to nonconforming data amounts to a violation of what is arguably the central principle of science, namely that when theory and data are in conflict we modify theory to accommodate data rather than modify data (as by under-representing the full range of relevant observations) to accommodate theory.

These remarks are not meant to be a criticism of any specific phonological theory. Rather, they are intended to question the assumption, common to most phonological theories, that all sound change must be phonetically motivated, along with the concomitant assumption that an appeal to natural classes in phonological change is impossible unless that change is driven by phonetic principles that arise from the physics of speech. The evidence considered here suggests instead that although the physics of speech may well draw speakers unconsciously into common types of sound change that are recurrent across the world's languages, conscious choice plays a role in various other sound changes, and in these changes speakers may make use of an innate knowledge of natural classes to produce phonological innovations of a purely conventional nature.

To conclude, it is generally agreed that most sound changes *are* phonetically motivated. However, this has never been the point at issue. Rather, the issue is whether this uncontroversial observation justifies an extension to the controversial claim that *all* sound changes are phonetically motivated. The only way that a certain answer will be obtained to this question is to catch an apparently saltatory change like Sundanese *w > *c-*, *-nc-* (Blust 2005) *in progress*. Proponents of the "Big Bang", in which the initial conditions of a sound change are invariably phonetic but may later be overlaid by socially-driven change, might argue that we have already seen a number of sound changes in progress, thanks to the pioneering work of William Labov over the past half century, and that this shows unequivocally that all sound change begins as phonetic process. But the reality is that the observation of sound change in progress has so far focused on a tiny fraction of the total number of changes that we must infer from reflexes in the world's languages, and no one can reasonably insist that this is an adequate sample for purposes of answering my question, "Must sound change be linguistically motivated?".

Acknowledgements

I wish to thank two anonymous reviewers for helpful suggestions with the presentation and particularly for drawing my attention to previously overlooked references, both of which led to improvements in an earlier version of this paper. Whatever shortcomings remain are mine alone.

References

Barnes, Robert H. 1974. *Kédang: A Study of the Collective Thought of an Eastern Indonesian People*. Oxford: Clarendon Press.

Beddor, Patrice Speeter, James D. Harnsberger & Stephanie Lindemann. 2002. Language-Specific Patterns of Vowel-to-Vowel Coarticulation: Acoustic Structures and Their Perceptual Correlates. *Journal of Phonetics* 30:4.591–627. doi: 10.1006/jpho.2002.0177

Beguš, Gašper. 2015. Intervocalic Devoicing in Kiput and Berawan Dialects. Paper presented at the 22nd Annual meeting of the Austronesian Formal Linguistics Association (AFLA), McGill University, Montreal, Canada, May 23, 2015.

Bender, Byron W. 1968. Vowel Dissimilation in Marshallese. *Working Papers in Linguistics* 1.1:88–96. Honolulu: Department of Linguistics, University of Hawai'i.

Blevins, Juliette. 2007. The Importance of Typology in Explaining Recurrent Sound Changes. *Linguistic Typology* 11:1.107–113. doi: 10.1515/LINGTY.2007.009

Blevins, Juliette. 2008. Phonetic Explanation Without Compromise: The Evolution of Mussau Syncope. *Diachronica* 25:1.1–19. doi: 10.1075/dia.25.1.02ble

Blevins, Juliette. 2009. Low Vowel Dissimilation Outside of Oceanic: The Case of Alamblak. *Oceanic Linguistics* 48:2.477–483. doi: 10.1353/ol.0.0045

Bloyd, Tobias. n.d. Reconstructing Proto-Sula-Mangon. University of Hawai'i qualifying paper.

Blust, Robert. 1970. *i* and *u* in the Austronesian Languages. *Working Papers in Linguistics* 2:6.113–145. Honolulu: Department of Linguistics, University of Hawai'i.

Blust, Robert. 1977. Sketches of the Morphology and Phonology of Bornean Languages: 1: Uma Juman (Kayan). *Papers in Borneo and Western Austronesian Linguistics* No. 2 7–122. (= *Pacific Linguistics*, A33.) Canberra: Pacific Linguistics.

Blust, Robert. 1980. Linguistic Evidence for Some Early Austronesian Taboos. *American Anthropologist* 83:2.285–319. doi: 10.1525/aa.1981.83.2.02a00020

Blust, Robert. 1988. Sketches of the Morphology and Phonology of Bornean Languages 2: Mukah Melanau. *Papers in Western Austronesian Linguistics*, No. 3 151–216. (= *Pacific Linguistics*, A78.) Canberra: Pacific Linguistics.

Blust, Robert. 1996. Low Vowel Dissimilation in Ere. *Oceanic Linguistics* 35:1.96–112. doi: 10.2307/3623032

Blust, Robert. 2000. Low Vowel Fronting in Northern Sarawak. *Oceanic Linguistics* 39:2.285–319. doi: 10.1353/ol.2000.0013

Blust, Robert. 2001a. Language, Dialect and Riotous Sound Change: The Case of Sa'ban. *Papers from the Ninth Annual Meeting of the Southeast Asian Linguistics Society*, 1999 ed. by Graham W. Thurgood, 249–359. Tempe: Arizona State University, Program for Southeast Asian Studies.

Blust, Robert. 2001b. Some Remarks on Stress, Syncope, and Gemination in Mussau. *Oceanic Linguistics* 40:1.143–150. doi: 10.1353/ol.2001.0002

Blust, Robert. 2004. *t to k: An Austronesian Sound Change Revisited. *Oceanic Linguistics* 43:2.365–410. doi: 10.1353/ol.2005.0001

Blust, Robert. 2005. Must Sound Change be Linguistically Motivated? *Diachronica* 22:2.219–269. doi: 10.1075/dia.22.2.02blu

Blust, Robert. 2007. Òma Lóngh Historical Phonology. *Oceanic Linguistics* 46:1.1–53. doi: 10.1353/ol.2007.0016

Blust, Robert. 2012. One Mark per Word? Patterns of Dissimilation in Austronesian and Australian Languages. *Phonology* 29:3.355–381. doi: 10.1017/S0952675712000206

Blust, Robert. 2013. *The Austronesian Languages (Second Revised Edition)*. (= *Asia-Pacific Linguistics Open Access Monographs*, A-PL 008.) Canberra: College of Asia and the Pacific, Australian National University.

Blust, Robert. 2016. *Austronesian Against the World: Where the P-map Ends*. Paper presented at BLS42, February 5–6, 2016.

Blust, Robert & Stephen Trussel. 2010–. *Austronesian Comparative Dictionary*. Available at http://www.trussel2.acd/com.

Brown, Herbert A. 1986. *A Comparative Dictionary of Orokolo, Gulf of Papua*. (= *Pacific Linguistics*, C84.) Canberra: Pacific Linguistics.

Burkhardt, Juergen. 2013. The Reconstruction of the Phonology of Proto-Berawan. Goethe University PhD dissertation.

Burkhardt, Juergen. 2015. How Did Long Terawan Berawan Develop Sixteen Vowel Phonemes? Paper presented at the Thirteenth International Conference on Austronesian Linguistics, Academia Sinica, Taipei, Taiwan, July 18–23, 2015.

Campbell, Lyle. 2013. *Historical Linguistics: An Introduction (Third Edition)*. Cambridge, Mass.: the MIT Press.

Clayre, Beatrice. 1996. An Introduction to Berawan Phonology. *The Sarawak Museum Journal* 71.209–240.

Cole, Jennifer, Gary Linebaugh & Bob McMurray. 2010. Unmasking the Acoustic Effects of Vowel-To-Vowel Coarticulation: A Statistical Modeling Approach. *Journal of Phonetics* 38:2.167–184. doi: 10.1016/j.wocn.2009.08.004

Comrie, Bernard. 1989. *Language Universals and Linguistic Typology (Second Revised Edition)*. Oxford: Blackwell, and Chicago: University of Chicago Press.

Dempwolff, Otto. 1934. *Vergleichende Lautlehre des austronesischen Wortschatzes 1: Induktiver Aufbau einer indonesischen Ursprache*. (= *Zeitschrift für Eingeborenen-Sprachen, Supplement 1*.) Berlin: Reimer.

Dierendonck, Martine van. n.d. *Kédang Vocabulary (Dutch-Indonesian-Kédang)*. Ms.

Donegan, Patricia. 1993. On the Phonetic Basis of Phonological Change. *Historical Linguistics: Problems and Perspectives* ed. by Charles Jones, 98–130. London and New York: Longman.

Fox, James J. n.d. *Lexicostatistical Lists for Languages of the Lesser Sunda Islands*. Ms.

Fromkin, Victoria & Robert Rodman. 1993. *An Introduction to Language (Fifth Edition)*. New York: Harcourt Brace Jovanovich.

García-Bellido, Paloma & Beatrice Clayre. 1997. Prosodic Constraints and Representations in the Berawan Word. *Oxford Working Papers in Linguistics, Philology and Phonetics* 2.17–52.

Garrett, Andrew. 2015. Sound Change. *The Routledge Handbook of Historical Linguistics* ed. by Claire Bowern and Bethwyn Evans, 227–248. London: Routledge.

Goddard, Ives. 2007. Phonetically Unmotivated Sound Change. *Verba Docenti: Studies in Historical and Indo-European Linguistics Presented to Jay H. Jasanoff by Students, Colleagues and Friends* ed. by Alan J. Nussbaum, 115–130. Ann Arbor: Beach Stave Press.

Gordon, Matthew. 1997. A Fortition-Based Approach to Balto-Fennic-Sámi Consonant Gradation. *Folia Linguistica Historica* 18.1–2:49–80.

Greenhill, Simon J., Robert Blust & Russell D. Gray. 2003–2017. *Austronesian Basic Vocabulary Database*. Available at https://abvd.shh.mpg.de/austronesian/.

Heston, Tyler. n.d. *A Reconstruction of Vowels and Stress in Proto-Timor-Alor-Pantar*. Ms., Honolulu: Department of Linguistics, University of Hawai'i.

Hoek, Michel van der. 2010. *Palatalization in West Germanic*. University of Minnesota PhD dissertation.

Horne, Elinor C. 1974. *Javanese-English Dictionary*. New Haven: Yale University Press.

Janda, Richard D. & Brian D. Joseph. 2003. Reconsidering the Canons of Sound-Change: Towards a 'Big Bang' Theory. *Historical Linguistics 2001: Selected Papers from the 15th International Conference on Historical Linguistics* ed. by Barry J. Blake, Kate Burridge & Jo Taylor, 205–219. Amsterdam/Philadelphia: John Benjamins. doi: 10.1075/cilt.237.14jan

Jong, J. P. B. de Josselin de. 1987. *Wetan Fieldnotes*. Verhandelingen van het Koninklijk Instituut voor Taal-, Land- en Volkenkunde 130. Dordrecht, Holland/Providence, USA: Foris Publications.

Keraf, Gregorius. 1978. *Morfologi dialek Lamalera*. Ende, Flores: Arnoldus.

Li, Paul Jen-kuei. 2004a [1981]. Reconstruction of Proto-Atayalic Phonology. In Li, ed., 625–692.

Li, Paul Jen-kuei. 2004b [1982]. Atayalic Final Voiced Stops. In Li, ed., 693–710.

Li, Paul Jen-kuei & Shigeru Tsuchida. 2006. *Kavalan Dictionary*. (= *Language and Linguistics Monograph Series*, A19.) Taipei: Institute of Linguistics, Academia Sinica.

Lichtenberk, Frantisek. 1994. Reconstructing Heterogeneity. *Oceanic Linguistics* 33:1.1–36. doi: 10.2307/3622999

Lynch, John. 2003. Low Vowel Dissimilation in Vanuatu Languages. *Oceanic Linguistics* 42:2.359–406. doi: 10.1353/ol.2003.0025

Nothofer, Bernd. 1981. *Dialektatlas von Zentral-Java*. Wiesbaden: Harrassowitz.

Odden, David. 2005. *Introducing Phonology*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511808869

Ohala, John J. 1981. The Listener as a Source of Sound Change. *Papers from the Parasession on Language and Behavior* ed. by Carrie S. Masek, Roberta A. Hendrick & Mary Frances Miller, 178–203. Chicago: Chicago Linguistic Society.

Posner, Rebecca. 1996. *Cambridge Language Surveys: The Romance Languages*. Cambridge University Press.

Ross, Malcolm. 2002. Mussau. *The Oceanic Languages* ed. by John Lynch, Malcolm Ross and Terry Crowley, 148–166. (= *Curzon Language Family Series*, 1.) Richmond, Surrey: Curzon Press.

Samely, Ursula. 1991. *Kedang (Eastern Indonesia), Some Aspects of its Grammar*. (=*Forum Phoneticum*, 46.) Hamburg: Helmut Buske.

Sneddon, James N. 1984. *Proto-Sangiric and the Sangiric Languages*. (= *Pacific Linguistics*, B91.) Canberra: Pacific Linguistics.

Soriente, Antonia. 2006. *Mencalèny & Usung Bayung Marang: A Collection of Kenyah Stories in the Òma Lóngh and Lebu' Kulit Languages*. Jakarta: Atma Jaya University Press.

Tadmor, Uri. 2003. Final /a/ Mutation: A Borrowed Areal Feature in Western Austronesian Languages. *Issues in Austronesian Historical Phonology* ed. by John Lynch, 15–35. (=*Pacific Linguistics*, 550.) Canberra: Pacific Linguistics.

Uhlenbeck, E. M. 1964. *Critical Survey of Studies on the Languages of Java and Madura*. (= *Koninklijk Instituut voor Taal-, Land- en Bolkenkunde, Bibliographical Series*, 7.) The Hague: Martinus Nijhoff. doi: 10.1007/978-94-011-8790-9

Author's address

Robert Blust
Department of Linguistics
University of Hawai'i at Manoa
1890 East-West Road, Moore Hall 569
Honolulu, HI 96822

blust@hawaii.edu