

# 1 *Irregular sound change and the post-velars in some Malakula languages*

---

JOHN LYNCH, University of the South Pacific

## 1. Introduction

There are a number of thorny little problem areas in Austronesian historical phonology, most of which have been investigated, at one time or another, by Bob Blust. In this volume in his honour, I would like to raise—or rather, re-raise—another thorny little problem, similar to one raised in Blust (1996).<sup>1</sup>

I will be examining the reflexes of just two protophonemes in just a handful of relatively closely related Oceanic languages, but the issue with which I am concerned is a considerably broader one: the fact that, despite the general principle that sound change is regular (or regular with certain definable exceptions), there are cases where this principle seems not to apply, and where a particular sound change does seem to be irregular.

The two protophonemes I will be concerned with here are the Proto Oceanic (POc) post-velars: the stop \*q and the trill \*R. Both of these have caused historical linguists considerable problems of one sort or another—in terms of what their position of articulation was, in terms of their manner of articulation (at least for \*R), and in terms of their widely varied reflexes—not least of which is the fact that they are probably lost more often than any other POc phonemes. I will concentrate on their reflexes in a number of Malakula languages for which we have reasonable amounts of data, most of which either have been only very recently published or are still unpublished.

I will be dealing in this paper with the reflexes of these two protophonemes in eleven languages spoken on the island of Malakula. The thirty or so actively spoken and moribund Malakula languages probably belong to a single sub-linkage within the Central Vanuatu linkage of the Southern Oceanic group, whose other members contain the non-Polynesian

---

<sup>1</sup> Bob Blust and I were graduate students together at the University of Hawai'i in the late 1960s and early 1970s. Although our paths have not physically crossed many times since then, I have always valued his comments on drafts of papers I have written, and have enjoyed reading his numerous and diverse contributions to Austronesian historical linguistics. I would also like to thank Andrew Pawley and an anonymous reviewer for comments on an earlier draft of this paper.

languages of the rest of Vanuatu and New Caledonia. There are probably two major subgroups of Malakula languages, though at this stage of research I cannot be definite about this, since there are a few languages which I cannot yet neatly classify. The languages I am dealing with here, together with their subgroup affiliation and the main sources of data, are as follows (within each subgroup I list languages roughly from north to south, and within examples I list languages in the order below):<sup>2</sup>

Eastern: Nese (Crowley 2006c)  
 Tirax (Amanda Brotchie, dictionary file)  
 Avava (Crowley 2006a)  
 Uripiv dialect of Northeast Malakula (Ross McKerras, dictionary file)  
 Unua (Elizabeth Pearce, dictionary file)

Western: V'ënen Taut (Fox 1979; Greg Fox, dictionary file)  
 Tape (Crowley 2006d)  
 Naman (Crowley 2006b)  
 Neverver (Julie Barbour, dictionary file)  
 Neve'ei (Musgrave 2001; Jill Musgrave and Terry Crowley, dictionary file)  
 Nāti (Crowley 1998)

## 2. Proto Oceanic \*R

Lynch, Ross & Crowley (2002:64) state that POc \*R “was probably a uvular trill, which is frequently lost or merged with a liquid (\*r or \*l) in daughter languages”. Wolff (2003:7), on the other hand, states that the ancestral phoneme in Proto Austronesian was “a voiced back spirant or possibly a back stop”, which he writes as \*ɣ rather than \*R and which he treats as the voiced equivalent of \*q. Given its reflexes in Oceanic languages—often a liquid or a fricative, and (almost?) never a uvular trill—it is possible that Wolff may be closer to the mark in assessing the phonetic nature of \*R.

Geraghty (1990:51) says that, “in the historical phonology and classification of Oceanic languages, probably no phoneme has been more extensively studied and used than \*R”. His very thorough study of reflexes of Proto Eastern Oceanic (PEOc) \*R shows that this protophoneme was definitely retained in all Vanuatu languages in some lexical items; where it is retained, it appears to merge with \*r. He postulated that “\*R is lost in proportion to distance from Western Oceanic, beginning in the Southeast Solomons” (1990:90). Based on whether \*R is retained or lost in particular lexical items, he suggested three major “boundaries” or isogloss bundles in Vanuatu: one between Mota and Raga, a second between Paama and Namakir, and a third between central Vanuatu and Erromango; this would divide the languages of Vanuatu into four groups: far north, north-central, Epi-Efate and southern.

Clark’s (in prep.) Proto North-Central Vanuatu (PNCV) reconstructions illustrate the nature of the general problem.<sup>3</sup> Examine the following:

<sup>2</sup> It goes without saying that I am extremely grateful to Julie Barbour, Amanda Brotchie, Greg Fox, Ross McKerras, Jill Musgrave and Elizabeth Pearce for so freely making available their unpublished data, and to the late Terry Crowley, who first got me interested in comparative Malakula linguistics.

<sup>3</sup> Clark in various publications has proposed that there is a North-Central Vanuatu (NCV) subgroup or linkage, and has made a large number of lexical reconstructions (Clark in prep.). While the integrity of this grouping is still under discussion within the wider context of the Southern Oceanic hypothesis (Lynch 2000), the reconstructions nevertheless have considerable validity.

- |     |          |                    |   |
|-----|----------|--------------------|---|
| (1) | POc/PEOc | PNCV               |   |
|     | a.       | *bakuRa<br>*Ropok  | *bakura ‘ <i>Calophyllum inophyllum</i> ’<br>*rovo ‘to fly’ |
|     | b.       | *cakaRu<br>*draRaq | *sakaRu ‘coral reef’<br>*daRa ‘blood’                       |
|     | c.       | *baReko<br>*Raka   | *baeko ‘breadfruit’<br>*aka ‘k.o. vine, <i>Pueraria</i> ’   |

Cases like those in (1a) show \*R merging with \*r and being reflected by the reflex of \*r in all or most of the languages which he considered. With those in (1b), on the other hand, \*R merges with \*r only in a few languages—usually Mota in Clark’s sample, and further investigation suggests that it is only the northernmost languages (Torres and Banks Is.) which retain it—but is lost in the remainder. Finally, cases like (1c) illustrate those where \*R is apparently universally lost in PNCV.

However, there is some counter-evidence to Geraghty’s hypothesis—cases where \*R is retained further south but apparently lost further north. Just a few examples are listed below:

(2)	POc/PEOc		*R LOST FURTHER NORTH	*R RETAINED FURTHER SOUTH
	*cakaRu	‘coral reef’	NE Ambae sakau, S Efate n/skau	Sye i/kri
	*cuRi-	‘bone’	Raga hui-, Paamese sī-	Sye no/ura-
	*Ruap	‘high tide’	Tamambo ua, Paamese ue	Lenakel e/lu-elu, Kwamera a/rə-ruk <sup>w</sup>

Let us now turn to the Malakula data to see what patterns can be identified in a lower-level subgroup of NCV with a large number of members. The first thing that needs to be said is that, in general terms, the retention or loss of \*R is fairly consistent across the languages I have been working with. That is to say, if \*R is retained in one of these languages in a particular lexical item, it is likely to be retained in all of them. Below are a few examples which illustrate this.<sup>4</sup>

---

<sup>4</sup> Reconstructions are POc unless preceded by N or C, indicating that they are attributed respectively to PNCV or PCV only. Blanks indicate no cognate or no data, while italics indicate that the form does not follow the general trend. V’ënen Taut and Nese have apicolabial consonants, written as *m̥*, *h̥*, etc. In Tirax, these apicolabials have changed further to alveolars: note \*m > n in *naxnal* < \*na-kamaliR ‘men’s house’. A number of languages have a prenasalised bilabial trill, written *B*.

**Table 1.** Regular retention of \*R

	*Rapi 'evening'	N *vaRa 'hand'	*kaRi(a) ' <i>Cordyline</i> '	*maRaŋo 'dry (coconut)'	*ma-wiRi 'left (side)'
Nese	revrav	na/vara-		naraŋ	
Tirax	revrev	vra-	kaɾe	mraŋ	maɛr
Avava	ki/drap	vara-	a/ari		m <sup>w</sup> iir
Uripiv	rivriv	ne/vre-	gari	raŋraŋ <sup>†</sup>	mair
Unua	revrev	vere-		mraŋ	yi/mair
V'änen Taut	kəna/raɣ	va-		məran	m̄ir
Tape	rivrip		na/arə/s	məraŋ	mor/ne-
Naman	revrev	nə/verə-	na/ɣari	meraŋ	ɣə/mir
Neverver <sup>‡</sup>	livrav	ne/vra-	na/xari	malləŋ	mer
Neve'ei	rivirav	ne/vera-	na/?ari	meraŋ	m <sup>w</sup> iyir
Nāti	revrev	ni/vara-	na/?ari		

<sup>†</sup> Uripiv unexpectedly loses initial \*ma- in this form.  
<sup>‡</sup> The occasional /l and /ll reflexes in Neverver are unexplained.

Similarly, if \*R is lost in one of these languages in a particular lexical item it is likely to be lost in all of them:

**Table 2.** Regular loss of \*R

	*Rum <sup>w</sup> aq 'house'	*kuRita 'octopus'	*tapuRiq 'conch'	*tuqaRi 'long ago'	*kamaliR 'men's house'
Nese	na/ine, n/em-	ne/ɣte	tavu	tua	na/ɣmal
Tirax	na/in		n/tav	tue	na/xnal
Avava	i/im	koit	a/taap	tua	amal
Uripiv	na/im	na/it	davō	tuwi	loln/amel
Unua	na/im	ɣuti/nbon	davu	tue 'forever'	ɣemer
V'änen Taut	nə/may	ɣut	na/tav	ti/tuei	n/amel
Tape	nə/may		tivwi	te/two	n/imel
Naman	ne/may	ni/ɣət		toɣe	na/amil
Neverver	a/iem	no/xoit	ne/tav	tue	na/xamal
Neve'ei	ni/yim	no/ɣoit	ne/tavu	tuyoi	na/?amal
Nāti		no/?oiyit	tāvu		na/?amel

These languages exhibit an overall consistency in their treatment of \*R—whether it is retained or lost—and there are only a handful of cases which show considerable inconsistency; these are illustrated in the table below, where retentions are in regular font and losses in italics:

**Table 3.** Irregular treatment of \*R

	N *[vo]m <sup>w</sup> aRaki 'ground dove'	*paRage ' <i>Pangium edule</i> '	*takuRu- 'back'	*suRuq 'juice, liquid'
Nese	<i>no/vomay</i>			<i>ne/jira-, nu/suwu-</i>
Tirax	<i>vemex</i>		<i>taxu-</i>	
Avava	<i>a/pmar</i>	<i>b<sup>w</sup>iki/var</i>	<i>u/dru-</i>	<i>a/sur, e/s-</i>
Uripiv	<i>wum<sup>w</sup>er</i>	<i>bik/wer</i>	<i>vitu</i>	<i>suwe-</i>
Unua	<i>na/mar</i>		<i>dure-</i>	<i>sue-</i>
V'änen Taut	<i>na/vimar</i>	<i>na/yak</i>	<i>du-</i>	<i>ui-</i>
Tape	<i>na/v<sup>w</sup>imar</i>	<i>nə/vək</i>	<i>e/taɣ</i> 'behind'	<i>jəre-</i>
Naman	<i>vomar</i>	<i>big/var</i>	<i>ne/tre-</i>	<i>ne/swe-</i>
Neverver		<i>nibig/var</i>		
Neve'ei	<i>ne/vimar</i>	<i>nabugu/var</i>	<i>ne/taʔa-</i>	<i>na/s-</i>
Nāti	<i>ne/vimar</i>	<i>ne/vaŋk</i>	<i>ni/taʔu-</i>	

With a few exceptions, then, \*R behaved fairly consistently in these languages (and indeed in PNCV), in the sense that it was retained in all languages in certain lexical items and lost in all languages in certain others. But is there any consistency in the patterns of loss and retention?

One general statement that can be made is that \*R was lost in absolute final position. There is one possible case of retention of root-final \*R when followed by possessive suffixes, but note that the reflexes show an additional vowel following the reflex of \*R:

- (3) \*ikuR- 'tail' > Uripiv *n/erure-*, Unua *ɣoɣore-*, Neverver *no/xore-*, Naman *no/ɣoɣorə-*

With the form \*maluR 'shade, shadow', final \*R seems to have been retained in Avava *milier* and Neve'ei *ni/milier*, but lost in Uripiv *na/mol*, Neverver *ni/milu-*, V'änen Taut *m̄ale/ka-* and Tape *m̄almilə-*. There is no evidence of final \*R being retained in any of these languages in the following etyma (though of course we do not have reflexes for every etymon in every language, either because the data are inadequate or the etymon was lost):

- (4) \*lasoR 'testicles'    \*qatoluR 'egg'    \*maturuR 'sleep'  
 \*qipaR 'in-law'    \*waiR 'water'    \*kamaliR 'men's house'  
 \*roŋoR 'hear'    \*saliR 'float'    \*madraR 'fermented breadfruit'  
 \*toŋoR 'mangrove'    \*niuR 'coconut'    \*pusuR 'bow and arrow'  
 \*sinaR 'shine'    \*rapu(R) 'ashes'

However, as Tables 1-3 show, there are cases of both retention and loss of both root-initial and root-medial \*R.<sup>5</sup>

Geraghty (1990:85) suggested that there may have been some phonological conditioning: e.g., initial \*R was likely to be lost before \*u but retained before \*a, medial \*R was more likely to be retained between identical vowels, etc. However, there does not seem to be any compelling evidence for this in the languages I have been examining, although there are a few trends. Table 5 summarises retention and loss of non-final \*R in all vocalic environments.

<sup>5</sup> There are virtually no cases of POC initial \*R occurring word-initially in these languages, since nouns generally have a fused article and verbs normally take one or more prefixes.

**Table 4.** Retention / loss of \*R in vocalic environments

V1↓	V2→	a	e	i	o	u	TOTAL
a		12 / 1		3 / 5	4 / 0	4 / 5	<b>23 / 11</b>
e		1 / 0		0 / 1		1 / 0	<b>2 / 1</b>
i		2 / 1		2 / 0			<b>4 / 1</b>
o		2 / 2	1 / 0		1 / 1	0 / 1	<b>4 / 4</b>
u		3 / 0		1 / 5		2 / 0	<b>6 / 5</b>
TOTAL		<b>20 / 4</b>	<b>1 / 0</b>	<b>6 / 11</b>	<b>5 / 1</b>	<b>7 / 6</b>	<b>39 / 22</b>

No clear-cut patterns emerge from the figures in Table 4. There are a couple of apparent tendencies, but none of these is exceptionless:

- there is a strong tendency for \*R to be retained before \*a and \*o;
- there is, as Geraghty suggested, a tendency for \*R to be retained between identical vowels (the figures from Table 4 are 17 retentions, 2 losses); and
- there is a tendency for \*R to be lost when before \*i (except / \*i \_\_ i).

However, there appear to be no particular patterns in relation to other vocalic environments; and, indeed, there are apparent “minimal pairs”, where \*R is retained in one item and lost in exactly the same vocalic environment in another item. For example, \*R is lost in the environment \*u \_\_ i in reflexes of \*tapuRiq and \*kuRita in Table 2, but retained in the following:

(5) \*tuRi ‘sew’ > Nese rur, Tirax drur, Avava tur, Uripiv o/tri, Neve’ei dur, Nāti tur

Similarly, \*R in the environment \*a \_\_ u is retained in (6a) below but lost in (6b):

- (6) a. \*yaRu ‘casuarina’ > Nese n/iar, Avava iar, Uripiv n/ur, Vēnen Taut ne/ier, Tape n/iar, Neve’ei n/iar, Nāti n/iar
- b. \*paRu ‘*Hibiscus tiliaceus*’ > Tirax na/ve, Uripiv vava, V’ēnen Taut yiyei, Tape vive, Naman nə/veve

Apart, then, from the almost universal loss of POc final \*R, it is not possible to define phonologically with any exactitude the conditions under which \*R is retained or lost. Its retention or loss appears on the surface to be quite random: it is retained in some lexical items, but lost in others.

### 3. Proto Oceanic \*q

When we examine the reflexes of \*q (apparently a uvular or back velar stop in POc) in these languages, we find a far greater degree of inconsistency than we do with \*R.

The only Vanuatu language which regularly retains \*q is Namakir, in which the reflex is the glottal stop (Sperlich 1989):

(7) POC		NAMAKIR	POC		NAMAKIR
*qasu	‘smoke’	ʔah	*toqa	‘fowl’	toʔ
*qusan	‘rain’	ʔih	*punuq	‘killed’	biniʔ
*paqoRu	‘new’	boʔo	*daRaqa	‘blood’	daʔ
*maqetom	‘dark’	maʔet	*taRaqa	‘cut’	daʔ

Recently, however, Lynch and Crowley (2003) pointed to occasional retentions of \*q in a few Malakula languages, and Lynch (2004) also noted occasional retention of \*q in

Southern Vanuatu languages. In this paper, I expand considerably on Lynch and Crowley's discussion.

### 3.1 The overt reflexes of \*q

POc \*q is lost far more often than it is retained. In this section I discuss the reflexes of \*q when it is not lost, without paying any attention to patterns of loss and retention, which I will discuss in §3.3. These reflexes are listed in Table 5, along with the reflexes of \*k for comparison; parentheses indicate a conditioned reflex, while the notation *i-i-* is shorthand for *i-*, *-i-* (i.e. *i* initially and medially).

**Table 5.** Overt reflexes of POc \*q and \*k in eleven Malakula languages

	*q-	*-q-	*-q	*k
Nese	-i-	-i-	∅	ʎ
Tirax	-i-		-x	x
Avava	i-, y-; -k-		-k-k	k (∅)
Uripiv	-i-		-i-	∅
Unua	ʎ-ʎ-		-ŋ-?; -g?	ʎ
V'ënen Taut	i-i-; -ʎ-	-ʎ	-ʎ-ʎ	ʎ (∅)
Tape	i-i-; w-w-	-w-; -ʎ-ʎ	-ʎ-ʎ	ʎ (∅)
Naman	i-i-; ʎ-ʎ-	-ʎ-	-ʎ-ʎ	ʎ (∅)
Neverver	x-x-; -i-	-i-	-x-x; (-k)	x
Neve'ei	?-?-; -i-; -ʎ-; w-w-	-ʎ-	-ʎ-ʎ; (-?-)	ʎ (?)
Nāti	-w-	-w-, -?	-?	?

Final \*q seems to have been retained in all of the eleven languages except Nese in at least one etymon, though in three of these (including the problematical Unua reflex of \*tob<sup>w</sup>aq below) only one etymon seems to show retention. In all cases except Uripiv (see \*tob<sup>w</sup>aq below) the reflex is a velar or glottal obstruent. Examples:

- (8) \*Rum<sup>w</sup>aq 'house' > V'ënen Taut nə/may, Tape nə/may, Naman ne/may  
 \*mimiq 'urinate' > Avava memek, V'ënen Taut məyei, Tape moy/wo, Neverver maxmax, Neve'ei maymay  
 \*tob<sup>w</sup>aq 'belly' > Tirax təbax, Uripiv depai-, Unua dabaŋo-, dobog, Naman dabaya-, Neverver ni/dəmxe-, Neve'ei ne/tabəʔa-  
 \*mataq 'new, raw' > Tirax mdrax, V'ënen Taut məday, Tape məday, Neverver mrex  
 \*luaq 'vomit' > Tape luay, Neverver lialuk, Neve'ei yoyyoʎ ?

Excluding the Uripiv and Unua cases, the obstruent reflex of \*q is the same as the reflex of \*k in all languages (except Nese, which has no obstruent reflex). Unua seems to show voicing crossover from \*k to \*g.

In non-final position the reflexes of \*q appear to be both a velar obstruent (or a glottal stop) and a high vowel/semivowel. Interestingly, medial \*q is rarely retained; but initial \*q is retained more frequently. When it occurred before a back vowel, \*q is normally reflected as *w*:

- (9) \*qone ‘sand’ > Tape nun/win, Naman dabano/wen, Neve’ei ne/wenwen, Nāti nempin/wen  
 \*qudu ‘palolo worm’ > Neve’ei nu/wud  
 \*quloc ‘maggot’ > Tape wiləs  
 \*qutin ‘penis’ > Neve’ei nu/wus-  
 \*maquRip ‘alive’ > Nāti mewur

but there are a couple of cases where some other reflex is found in some languages:

- (10) \*q > w other reflex  
 \*quluŋ-an ‘pillow’ > Neve’ei wulwul Avava u/kulaŋ, V’ënen Taut n/iululna- (v.)  
 \*qusan ‘rain’ > Nāti nu/wuh (n.) Tape iu (v.), Naman ius (v.)  
 \*qutan ‘inland’ > Nāti ne/wut Unua ve/γut, Naman γaut, Neve’ei ?out  
 \*leqo- ‘voice’ > Nāti ni/lo?

When not before back vowel, \*q is sometimes retained as a velar or glottal:

- (11) \*qalawa- ‘sibling’s child’ > Neverver xala- ‘nephew, uncle’  
 \*qaloŋo ‘*Acanthurus* sp.’ > Unua va/γaro  
 \*qapaRa- ‘shoulder’ > Naman no/γoverə- ‘wing’, Neverver na/xarevra- ‘wing’  
 \*qatoluR ‘egg’ > Unua γori- V’ënen Taut na/γadrəl, Neve’ei na/?adle-  
 \*laqia ‘ginger’ > Tape ləyləy  
 \*tuqaRi ‘long ago’ > Naman toye, Neve’ei tuyoi, duyoi

In other cases, the reflex is *i*. I give the data first (and see also \*quluŋ-an and \*qusan in (10) above), and then discuss it.

- (12) \*qaŋaRi ‘*Canarium*’ > Avava yaŋa, Uripiv n/iŋi, Tape n/ieŋe, Neverver n/iŋa, Neve’ei n/iŋi  
 \*qaRa(r) ‘fence’ > Nese, Tirax, Neverver, Neve’ei n/iar  
 \*qasu ‘smoke’ > Nese n/ies, V’ënen Taut ie-naŋ, Naman ies (v.), n/iisə- (n.)  
 \*qatop ‘sago’ > Nese, Tirax, Uripiv, Neverver, Neve’ei n/iat, Avava iat, V’ënen Taut ne/iet; Tape, Naman n/iet

Generally, there is consistency between these languages as to whether a velar or a high vowel is the reflex (though not as to whether \*q is retained or not). The one significant case of inconsistency is the following:

- (13) \*qase- ‘jaw, chin’ > **velar:** V’ënen Taut na/γ-, Neverver na/xas-, Neve’ei na/γase-  
**i:** Nese n/ias-, Uripiv n/ise-, Tape n/isi-

Now *i* occurs as a putative reflex of \*q mainly when \*q was initial and mainly in nouns, and thus it often occurred preceded by an article. The comparison \*na-qatop ‘sago’ >



V'ënen Taut *ne/iet*, however, is the only case where we can be absolutely sure that \*q became *i*; that is:

*	n	a	-	q	a	t	o	p
	↓	↓		↓	↓	↓		
	n	e		i	e	t		

In all other cases, we need to assume that \*na-qa.. > *na-ia.* > *nia.* (sometimes further > *nie.* or just *ni.*). However, there is some evidence from Avava which suggests that this is the correct interpretation. Avava reflects the POC article \*a, not \*na, and this often underwent vowel harmony (thus \*bokasi ‘pig’ > a/buah, \*mwata ‘snake’ > a/mwat, \*toqa ‘fowl’ > o/to, \*paRi ‘stingray’ > e/ve, \*mwele ‘cycad’ > i/mwil, etc.). Noun-initial a can reflect both the article (when retained) or the first vowel of the root (when the article was not fused),<sup>6</sup> and note that in the latter case there is no prothetic consonant (as there is with a number of languages—e.g. Fijian underlying word-initial a takes a prothetic y):

(14)	POC	FUSED ARTICLE	POC	ROOT-INITIAL a		
	*a-baga	‘banyan’	a/baŋ	*kamaliR	‘men’s house’	amal
	*a-manuk	‘bird’	a/man	*katabola	‘dragon plum’	atipol
	*a-niuR	‘coconut’	a/ni	*kavika	‘Malay apple’	avik

It will be seen from the data in the right hand column that root-initial a generally occurs when initial \*k has been lost. In cases where this happens and the article is retained, it coalesces with root-initial a as a long vowel:

(15)	POC		AVAVA
	*a-kabu	‘fire’	a/aB
	*a-karia	‘cordyline’	a/ari

There are only two *ia*-initial nouns in the Avava lexicon: *iar* ‘casuarina’ < \*yaRu and *iat* ‘sago’ < \*qatop, but there is also *yaŋa* ‘canarium’ < \*qaŋaRi. I suggest that in the latter two cases the *i* and *y* clearly reflect \*q, and that \*a-qatop > *a-iatop* > *iat*, while *yaŋa*, being bi-moraic, did not take the article. That is, there is no evidence of prothetic *i* or *y* in Avava, and there is evidence that \*q > *i* or *y*.

Although we can predict that final \*q will merge with \*k as a velar obstruent and that *w* is the usual reflex of \*q before a back vowel, it is more difficult to predict when non-final \*q is reflected as a velar and when it is reflected as *i* or *y*, since both are found in similar environments (particularly \*na- \_\_ a).

### 3.2 The reflexes themselves

The discussion above has shown that, when retained, final \*q is normally reflected as a velar or glottal obstruent, while non-final \*q is reflected as *w* before \*u and as *i* or the corresponding glide or else a velar elsewhere. Although not strictly germane to the overall topic of sporadic reflexes, it is worth briefly discussing why these different reflexes may have developed.

---

<sup>6</sup> Avava is one of a number of Malakula languages in which the article is fused if the root contains one mora but is normally not retained if the root is bi- or multi-moraic (see Lynch 2007) – though \*a-karia > *a/ari* in (15) seems to be an exception.

Final \*q merges with \*k in all of the eleven languages which reflect it. Given that POc \*q was close to \*k in articulatory terms, there is nothing very unusual about it having a velar reflex.

It is unusual, however, for \*q to have a high vowel/glide reflex, which is what happens with some cases of retained non-final \*q. First, though, this clearly shows that \*q in this environment did not merge with \*k; and so we can assume that \*q was retained as a distinct phoneme in any “Proto Malakula”. I assume that \*q eventually became a glide in this environment, later vocalising in some environments: a direct \*q > *i* change seems highly unlikely.

There is, however, some unpredictability as to what the non-final reflex is. Some languages show only a velar (Unua), some only a high vowel or glide (Nese, Uripiv, Tirax), but others show both, and there seems to be no phonological conditioning involved. Indeed, considering \*na-qa-initial forms, \*q surfaces at least once as *i* and at least once as a velar in Neverver, V’ënen Taut, Naman and Neve’ei.

One possible explanation may be that \*q became a fricative in this environment, which was basically intervocalic. This fricative may have been uvular or velar and, given the intervocalic environment, it may have been voiced rather than voiceless: i.e. [ɣ] or [y]. (This must have predated any lenition of \*k in this environment, though.) If this was the case, we would then need to assume [ɣ] > [w] before back vowels and sometimes > [y], sometimes to (or remaining) a velar before non-back vowels. Both seem to be more natural developments.

### 3.3 Loss and retention of \*q

Having established the reflexes of \*q when it is retained, I move now to look at overall patterns of retention and loss. The first thing that needs to be said is that the situation is somewhat different from that of \*R: even in cases where some languages show retention of \*q, others show loss in the same etymon. Table 6 illustrates this: retentions are in regular font, losses in italics.

**Table 6.** Irregular treatment of \*q

	*qase- ‘chin, jaw’	N *qaŋaRi ‘ <i>Canarium indicum</i> ’	*tuqaRi ‘long ago’	*mimiq ‘urinate’
Nese	n/ias-	<i>n/eŋa</i>	<i>tua</i>	
Tirax	<i>n/he-</i>	<i>n/eŋa</i>	<i>tue</i>	<i>muŋe?</i>
Avava	as-	yaŋa	<i>tua</i>	memek
Uripiv	n/ise-	n/iŋi	<i>tuwi</i>	<i>meme</i>
Unua	<i>n/ese-</i>	<i>n/eŋe</i>	<i>tue</i> ‘forever’	<i>me, meme</i>
V’ënen Taut	na/y-		<i>ti/tuei</i>	məyei
Tape	n/isi	n/ieŋe	<i>te/two</i>	moywo
Naman	<i>na/ase-</i>	<i>n/eŋe</i>	toye	<i>mimi</i>
Neverver	na/xas-	n/iŋa	<i>tue</i>	maxmax
Neve’ei	na/yase-	n/iŋi	tuyoi, duyoi	maymay
Nāti	<i>n/ehe-</i>	<i>n/eŋei</i>		<i>mimim</i>

POc \*q has no overt reflex in any of these languages in around 60 etyma. Figures for its retention (excluding doubtful cases) in each of the eleven languages are as follows:

**Table 7.** Retention of \*q

	*q-	*-q-	*-q	TOTAL		*q-	*-q-	*-q	TOTAL
Nese	4	1		<b>5</b>	V'ënen Taut	3	3	5	<b>11</b>
Tirax	2		1	<b>3</b>	Tape	9	4	5	<b>18</b>
Avava	3		2	<b>5</b>	Naman	6	1	2	<b>9</b>
Uripiv	3			<b>3</b>	Neverver	6	1	5	<b>12</b>
Unua	3		1	<b>4</b>	Neve'ei	10	1	5	<b>16</b>
					Nāti	5	3	1	<b>9</b>

It is clear from Table 7 that \*q is retained in only a minority of cases. It is also clear that retention figures are higher among languages of the putative Western subgroup (on the right of the table above) than those of the Eastern subgroup (on the left).<sup>7</sup>

I do not need to list many examples of loss of \*q. The following handful shows no retention of \*q in four phonological environments in which it is retained in some items in some languages (see 3.1), showing that there seems to be no phonological conditioning involved.

**Table 8.** Loss of \*q

	*qupi 'yam'	C *kumaqu 'Intsia bijuga'	*suRuq 'fluid, juice'	*puaq 'fruit'
Nese			ne/jira-, nu/suwu-	no/vo-, nu/vu-
Tirax				bi/vu
Avava	o/ovi 'k. yam'		a/sur	va/na-
Uripiv	n/ov 'k. yam'		suwe-	we/ne-
Unua				ve/ne-
V'ënen Taut		ɣəmau	ui-	na/va-
Tape		nə/ɣmo	jəre-	no/vo-
Naman		no/ɣmo	ne/swe-	na/va-
Neverver		nu/xuma		
Neve'ei	n/obi 'k. yam'	nu/ɣumo	na/s-	ne/v <sup>w</sup> e-
Nāti		ne/?umou		nö/van

To summarise: \*q was normally lost, but was sometimes retained. Although we can make some predictions about what its reflex is when it is retained, we can make no prediction on any phonological basis about whether or not it is retained (a) in a particular etymon or phonological environment and (b) in a particular language—apart from just possibly suggesting that \*q is more likely to be retained than lost initially before \*a. (Even two such closely related languages as V'ënen Taut and Tape are inconsistent in their retention of \*q in particular items.)

#### 4. Summary of Malakula data

To summarise the data on these two protophonemes in these languages:

<sup>7</sup> The amount of data available varies from language to language, but the disproportionate number of retentions of \*q in the west has nothing to do with that. We have, for example, more data on Uripiv than any of the others, and yet Uripiv has just about the smallest number of retentions; we have more data on Tirax, Unua and Avava than on Nāti, and yet the latter shows at least twice as many retentions.

- with \*R, there is consistency as among the various languages as to whether or not \*R was retained in any particular lexical item;
- with \*q, there is no consistency as among the various languages as to whether or not \*q was retained in any particular lexical item;
- with both, there is basically no consistency in terms of phonological environment as to when each was retained and when it was lost, apart from one or two minor trends; and
- with \*q there is no consistency as to whether its non-final reflex is a glide or a velar obstruent.

The behaviour of \*R in these language I have been looking at is not dissimilar to what happens to \*R elsewhere in Vanuatu, so to that extent Malakula is a microcosm of the rest of Vanuatu. However, the behaviour of \*q is quite different: \*q is regularly retained in Namakir (and nowhere else in Vanuatu) and in parts of New Caledonia; elsewhere in this area it seems to be regularly lost (with the exception of the very few apparent retentions in Southern Vanuatu described in Lynch 2004). The Malakula case, then, is *not* a microcosm of the rest of this area as far as \*q is concerned.

The regular retention of \*q in Namakir and New Caledonia implies that \*q was regularly retained in Proto Southern Oceanic and its various daughter-languages (unlike \*R, which was irregularly lost). The total loss of \*q in many of the descendants of Proto Southern Oceanic thus has to be seen as having occurred through a number of independent but probably identical changes. However, at least as far as Malakula languages are concerned, \*R and \*q are similar in that they are sometimes retained and sometimes lost.

## 5. Irregular sound change

One of the tenets of historical linguistics is that sound change is regular. This dates back to the latter decades of the nineteenth century, when the Neogrammarians (*Junggrammatiker*) declared that sound “laws” were without exception. Since then, the view has moderated somewhat, but it can be fairly said that most historical linguists would hold that sound changes are *generally* regular, and that while there may be exceptions all or most of these can be explained in some way.

### 5.1 Exceptions to regular sound change

Three of the explanations often put forward—quite legitimately—to explain exceptions are analogy, avoidance of homophony and contact between related languages. Koch (1996:220), for example, shows the operation of analogy to produce a phonological irregularity in Semitic languages. Note first the following verb inflections:

(16)		1SG	2SG
	Proto Semitic	*-ku	*-ta
	Akkadian	-ku	-ta
	Arabic	-tu	-ta
	Ethiopic	-ku	-ka

Akkadian is regular, but Arabic and Ethiopic “are each assumed to have altered their consonant by analogy with the other member of their respective paradigms”: Arabic shows irregular \*k > t in 1SG, Ethiopic irregular \*t > k in 2SG.

Campbell (1996:77-78) notes how avoidance of homophony may bring about lexical replacements and phonological irregularity, and gives as one example of the latter certain German dialects in which two regular changes—loss of intervocalic *g* and unrounding of *ü*—would have meant that *liegen* ‘lie (down)’ and *lügen* ‘lie (= tell lies)’ would be homophonous. Instead, these changes do not apply in just these two words, to avoid this homophony.

There are numerous cases within Oceanic of contact between related languages producing what appear to be irregular phonological changes: indeed, this irregularity is one of the techniques for identifying borrowings. For example, in non-Polynesian languages of Southern Vanuatu POC final vowels are regularly lost: \**rua* ‘two’ > Lenakel *k/iu*, \**tolu* ‘three’ > *kə/sil*, \**kani* ‘eat’ > *kən*, etc. Cases of irregular retention of final vowels do occur: \**kiajo* ‘outrigger boom’ > *nə/kiatu*, \**jila* ‘sail’ > *tila*, etc. But it is apparent that these words have been borrowed from the nearby Polynesian Outlier language Futuna-Aniwa, which retains POC final vowels.

However, while most sound changes may well be perfectly regular, or may be almost perfectly regular but have some explainable exceptions, this is not always the case. For example, Blust (1996:137) says that:

To a large extent the success of the Neogrammarian hypothesis has stemmed from the apparent fact that sound change *is* overwhelmingly regular. Where irregularities exist it has generally been found possible to explain them (or, all too often, to ‘explain them away’) as products of borrowing, analogy, or some other mechanism of secondary change. At the same time the problems associated with the Neogrammarian hypothesis stem from two apparent facts which may conflict with it: (1) The regularity of phonological change is an epiphenomenon rather than a primary datum, and (2) despite its overwhelming regularity, not *all* sound change is regular.

And again:

Lest I be misread, let me emphasize in the strongest terms that I do not advocate a facile acceptance of irregularity in sound change. Every effort should be made to find rule-governed explanations for the primary observations. But when plausible explanations for irregularity cannot be stated it is pointless to resort to mechanical contrivances out of fear that the only alternative to such ad hoc solutions is to open a Pandora’s box of methodological chaos. Irregularity is not mere chaos. Rather, ... irregularity appears to be an integral part of the natural process of language change (Blust 1996:153).

A number of studies in Durie and Ross (1996) examine real or apparent irregularity in phonological development. In their introduction, the authors discuss lexical diffusion of phonological changes in a number of the case studies in the following terms:

[A]t least in some of these cases, a lexically diffusing sound change has been halted before it completed its journey through the lexicon. ... [A] speaker-oriented version of the hypothesis can be formulated. It says:

- i. that each speaker who adopts a sound change does so first as part of the orderly variation of that speaker’s speech;
- ii. that this variation progressively shifts in favour of the ‘new’ sound;

- iii. that (ii) applies initially only to certain items in the lexicon which contain the relevant sound in the relevant environment, then progressively on to other elements (Durie and Ross 1996:23).

The assumption is that with a wholly regular change, process (iii) continues to apply until all appropriate items are affected for all speakers who have undergone the change. With an irregular change, however, process (iii) stops before applying to all appropriate items.

## 5.2 POc \*R and \*q in Malakula

Can the sporadic loss of \*R and \*q in Malakula be explained by factors like analogy, avoidance of homophony or contact with related languages? Or do these changes fit the truly irregular case? And if the latter, why?

I cannot see that analogy has any role here, since this usually occurs in cases like the Semitic one above, where paradigm sets are involved; in the Malakula cases I have been talking about, we are dealing with consonants which are root-initial, -medial and -final, and affixal morphology does not enter into it.

Avoidance of homophony also does not seem to be an issue here. Let us take the case of \*R. We can make one of two assumptions:

- a. The regular change was \*R > *r*, but \*R was lost to avoid homophonous forms with *r* from other sources (mainly \*r). As an imaginary example, assume a form \*maRu ‘dog’ which should become *maru*, but because there is already \*maru ‘snake’ > *maru* the former loses \*R to avoid this homophony and becomes *mau*.
- b. Conversely, the regular change was \*R >  $\emptyset$ , but this was blocked when a homophonous form would have been produced, and \*R remained *r*. As another imaginary example, assume a form \*baRi ‘tree’ which should become *pai*, but because there is already \*bai ‘fruit’ > *pai* the former retains \*R to avoid homophony (\*baRi becoming *pari*).

This is the kind of change which might account for one or two—even perhaps half a dozen—exceptions (as in the German case above), in perhaps one or two languages. But I cannot see this happening on such a large scale as we have with reflexes of \*R and, to a slightly lesser extent, \*q in Malakula languages—a large scale in terms of both the number of etyma and the number of languages involved.

Contact with neighbouring and related languages was probably frequent in Malakula, given the small size of most of the language-communities there, and borrowing from one language to another undoubtedly took place. Given the situation described for \*q above, it *might* be just possible to imagine that some language(s) which regularly lost \*q borrowed the occasional etymon from some other neighbouring language in which \*q was retained (or vice versa), even though in making this assumption we would have to allow for rather widespread borrowing of “basic” vocabulary (and for the fact that many cognate morphemes are formally different in other respects as well). But this situation is just unimaginable with \*R, given the regularity of its loss or retention as between different languages. That is:

- we would have to assume that in a number of languages in which \*R >  $\emptyset$  was regular, *the same lexical items* were borrowed from certain other languages in which \*R > *r* was regular; and/or

- we would have to assume that in a number of languages in which \*R > r was regular, *the same lexical items* were borrowed from certain other languages in which \*R > ∅ was regular.

These assumptions are untenable.

I think the only conclusion that can be drawn is that we are dealing here with a true case of irregularity. In Malakula (and many other Vanuatu languages as well), \*R—which must have been phonetically similar to \*r—began to be lost finally and probably before high vowels. This change, however, was not completed before a second change took place: the merger of \*R and \*r. Thus \*R is lost in some lexical items but retained as the reflex of \*r in other items in the same phonological environment. Something similar happened with \*q, although here the process of lenition and subsequent loss was further advanced before the remaining reflexes of \*q merged with a velar in some environments and a high vowel (via a glide) in others.

Both cases which I have detailed in this paper, then, illustrate the fact that *some* sound changes can be truly irregular.

## References

- Blust, Robert. 1996. The Neogrammarian hypothesis and pandemic irregularity. In Durie and Ross (eds), 135-156.
- Campbell, Lyle. 1996. On sound change and challenges to regularity. In Durie and Ross (eds), 72-89.
- Clark, Ross. In prep. *North and Central Vanuatu: A comparative study*.
- Crowley, Terry. 1998. A salvage sketch of Nāti (Southwest Malakula, Vanuatu). In D.T. Tryon (ed.) *Papers in Austronesian linguistics No. 5*, 101-148. Canberra: Pacific Linguistics.
- \_\_\_\_\_. 2006a. *The Avava language of Central Malakula* [ed. by John Lynch]. Canberra: Pacific Linguistics.
- \_\_\_\_\_. 2006b. *Naman: A vanishing language of Malakula (Vanuatu)* [ed. by John Lynch]. Canberra: Pacific Linguistics.
- \_\_\_\_\_. 2006c. *Nese: A diminishing speech variety of Northwest Malakula* [ed. by John Lynch]. Canberra: Pacific Linguistics.
- \_\_\_\_\_. 2006d. *Tape: A declining language of Malakula (Vanuatu)* [ed. by John Lynch]. Canberra: Pacific Linguistics.
- Durie, Mark, and Malcolm Ross. 1996. *The comparative method reviewed: Regularity and irregularity in language change*. New York: Oxford University Press.
- Fox, G.J. 1979. *Big Nambas grammar*. Canberra: Pacific Linguistics.
- Geraghty, Paul. 1990. Proto-Eastern Oceanic \*R and its reflexes. In J.H.C.S. Davidson (ed.) *Pacific Island languages: Essays in honour of G.B. Milner*, 51-93. London and Honolulu: School of Oriental and African Studies and University of Hawaii Press.
- Koch, Harold. 1996. Reconstruction in morphology. In Durie and Ross (eds), 218-263.

- Lynch, John. 2000. Linguistic subgrouping in Vanuatu and New Caledonia. In Bill Palmer & Paul Geraghty (eds), *Proceedings of the Second International Conference on Oceanic Linguistics (SICOL), Vol. 2: Historical and descriptive studies*. Canberra: Pacific Linguistics, 505:155-184.
- \_\_\_\_\_. 2004. Bilabial reflexes of Proto-Oceanic \*q in Southern Vanuatu languages. *Oceanic Linguistics* 43,1:214-220.
- \_\_\_\_\_. 2007. The Malakula linkage of Central Vanuatu. Paper presented to the Seventh International Conference on Oceanic Linguistics, Nouméa, New Caledonia, July 2007.
- Lynch, John, and Terry Crowley. 2003. More reflexes of Proto-Oceanic \*q in Vanuatu languages. *Oceanic Linguistics* 42,1:233-238.
- Lynch, John, Malcolm Ross and Terry Crowley 2002. *The Oceanic languages*. London: Curzon Press.
- Musgrave, Jill. 2001. *A grammar of Nefe'ei, Vanuatu*. Unpublished MPhil thesis, University of Waikato.
- Sperlich, Wolfgang B. 1989. Is the Na-makir (Central Vanuatu) glottal stop a reliable witness for POC \*q? In Ray Harlow and Robin Hooper (eds) *VICAL 1: Oceanic Languages (Papers from the Fifth International Conference on Austronesian Linguistics)*, 527–553. Auckland: Linguistic Society of New Zealand.
- Wolff, John U. 2003. The sounds of Proto Austronesian. In John Lynch (ed.) *Issues in Austronesian historical phonology*, 1-14. Canberra: Pacific Linguistics.