Where *R they all? The Geography and History of *R-loss in Southern Oceanic Languages

Alexandre François

Where *R they all? The Geography and History of *R-loss in Southern Oceanic Languages

Alexandre François

LANGUES ET CIVILISATIONS À TRADITION ORALE (CNRS), PARIS, AND AUSTRALIAN NATIONAL UNIVERSITY

Some twenty years ago, Paul Geraghty offered a large-scale survey of the retention and loss of Proto-Oceanic *R across Eastern Oceanic languages, and concluded that *R was "lost in proportion to distance from Western Oceanic." This paper aims at testing Geraghty's hypothesis based on a larger body of data now available, with a primary focus on a tightly knit set of languages spoken in Vanuatu. By observing the dialectology of individual lexical items in this region, I show that the boundaries between languages retaining vs. losing *R differ for each word, yet they all define a consistent north-to-south cline whereby *R is lost in the south. This cline, which confirms Geraghty's observations, can be recognized all the way to southern Vanuatu and New Caledonia. Such a neat geographic distribution observed today can be interpreted in historical terms. I propose that the tendency to lose *R emerged somewhere south of Efate, at an early date in the settlement of the archipelago. This sound change triggered a range of individual lexical innovations, each of which spread across what was then a vast social and linguistic network, encompassing the whole of Vanuatu and New Caledonia. The geography of *R reflexes constitutes a fossilized picture of prehistoric social networks, as the once unitary world of Lapita settlers was beginning to break down into increasingly diversified dialects—the ancestors of modern languages.

1. PRESENTATION. This study surveys the languages of Vanuatu—and to a lesser extent, of New Caledonia—in order to observe the geographical patterns in the retention and loss of a particular consonant, namely Proto-Oceanic (POC) *R.[1]

1. I wish to thank LACITO (Langues et Civilisations à Tradition Orale), the Centre National de la Recherche Scientifique, and the French Ministère de la Recherche for funding my field trips to Vanuatu since 2003; as well as the Linguistics Department at the Australian National University for its intellectual support during the writing of this paper. I am also grateful to Andrew Pawley, Malcolm Ross, and especially John Lynch and Paul Geraghty, for their helpful comments on earlier versions of this article; any remaining errors are mine. My gratitude also goes to several people who provided firsthand data from various languages: Julie Barbour for Neverver, Isabelle Bril for Zuanga, Ross Clark for Northern Santo, Laura Dimock for Nahavaq, Gaïa Fisher for Sungwadaga, Murray Garde for Sa, Valérie Guérin for Maŭea, Agnès Henri for Sungwadia, Sébastien Lacrampe for Lelepa, and Cindy Schneider for Abma. Finally, such a study would not have been possible without the help and kindness of the many people from northern Vanuatu who patiently taught me their languages.

Oceanic Linguistics, Volume 50, no. 1 (June 2011)

In this introduction, I begin in 1.1 by recapitulating what we know of the phonetics of POc *R. Then I summarize the initial findings in Geraghty (1990) regarding the patterns of *R-loss in the Pacific, and more specifically for North-Central Vanuatu languages (1.2). This area is the focus of the next section, section 2, in which I bring new data to the examination of the patterns of retention and loss of *R in individual words. In section 3, I will widen the scope of observation and comment on the distribution of *R throughout the whole Southern Oceanic linkage—that is, Vanuatu and New Caledonia.

After establishing the spatial distribution of *R reflexes in this region, I shall discuss the implications of these empirical results regarding the linguistic and social history of the region in sections 4 and 5. My hypothesis will be that the modern dialectology of *R reflects the spread of an irregular sound change by lexical diffusion across a vast dialect chain, shortly after the initial settlement of Island Melanesia by Lapita colonizers. This process took the form of several individual lexical innovations that spread northward across what was then a vast chain of mutually intelligible dialects, forming a unified social network.

1.1 THE PROTOPHONEME *R. Among the consonants that can be reconstructed in study of the history of Oceanic languages, the phoneme *R is probably the one that has received most attention in the literature over recent decades, from Milke (1958) to Lynch (2009a). The most thorough study of this consonant is undoubtedly Geraghty (1990), itself based on Geraghty (1978).

This consonant raises a number of issues for the historical linguist. First of all, there is uncertainty about its phonetic nature, whether in Proto-Austronesian (PAN) or in later interstages, such as Proto-Oceanic. Examining a number of Austronesian (mainly non-Oceanic) languages, Blust (2009:582) lists as many as twenty different reflexes of PAN *R, ranging from /g/ or /n/ to /s/ or /y/, not to mention its frequent loss. Reflexes of *R within the Oceanic subgroup are less exuberant: the typical situation is that *R merges with *l in some languages, with *r in others (Ross 1988:31, Ross and Næss 2007:472), or disappears altogether.

Ross (1998:16), writing about POc, suggests that "probably the phoneme *r was an alveolar trill … whilst *R was probably a uvular trill, easily lost or merged with *r or *l in daughter languages." The problem with this hypothesis is that, in modern Oceanic languages, *R is reflected nowhere as a uvular rhotic; besides, we would have to explain how so many languages could have changed a uvular [ʀ] into an alveolar trill [r] (or a lateral [l] for that matter)—a change that is hardly attested anywhere in the world, unlike its symmetrical counterpart *[r] > [ʀ]. More promising is the suggestion by Blust (2009:582) that both *r and *R were alveolar rhotics in Proto-Austronesian, *R being a trill and *r a flap. It is not uncommon in the world for a language to contrast a flap and a trill,[2] yet one can easily imagine cases of merger or phonetic instability of such a contrast. The opposite distribution to the one proposed by Blust (*r a trill, *R a flap) probably deserves to be con-

2. A well-known example is Spanish. The contrast /r/ ≠ /r/ is rarer among Oceanic languages, yet is attested, for example, in Araki (François 2002:18). Note, however, that the Araki pair does not reflect directly a former pair *r vs. *R: Araki /r/ reflects *t, whereas /r/ reflects both *r and *R (see 2.2.1). Kairiru, a Western Oceanic language of Papua New Guinea (Wivell 1981), also has /r/ ≠ /r/, but the precise relationship between these two phonemes and reconstructed *r, *R, and *dr is complex (Malcolm Ross, pers. comm.).

sidered too, especially if *R is attested to merge with an alveolar lateral *l. The final answer to this question is not crucial to the present study. Furthermore, it will focus on Southern Oceanic, an area where the segmental reflexes of *R hardly ever contrast with those of *r. *R can henceforth be understood as a liquid, with no further specification.

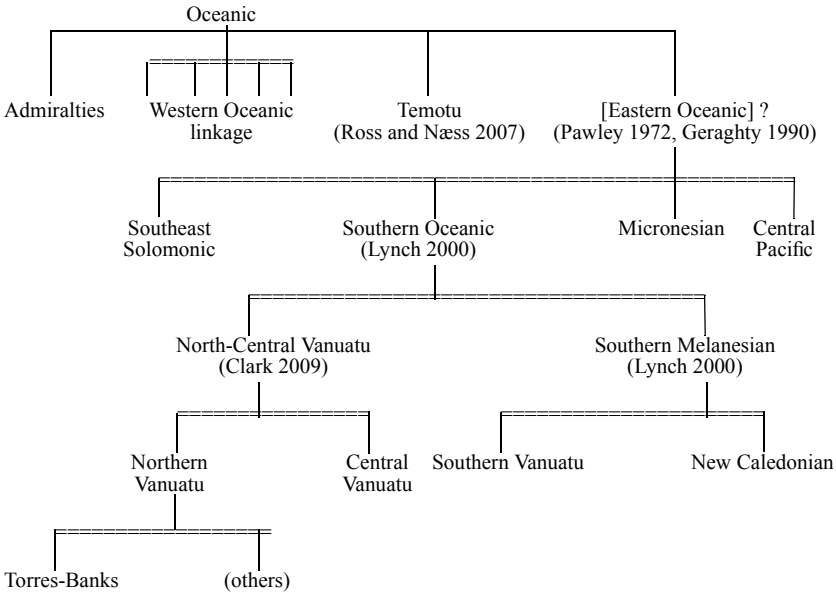
1.2 LOSS AND RETENTION OF *R: STATE OF THE ART. More essential to the present paper is another property of *R, namely its tendency to be lost in ways that may be significant to the historical linguist.

Geraghty (1990) observed how *R is retained or lost in the putative subgroup known as Eastern Oceanic—that is, a subgroup of Oceanic that he understood at that time as comprising Southeast Solomonic (SES) languages, North-Central Vanuatu (NCV),[3] Micronesian, and Central Pacific (Fijian, Rotuman, and Polynesian).[4] The classification has evolved since then in various respects (see Ross, Pawley, and Osmond forthcoming). I shall only mention two of these changes that are relevant to the present study. First, "Eastern Oceanic," first proposed by Pawley (1972) and used by Geraghty (1990), has since then come into question as a well-established node (Ross, Pawley, and Osmond 2008:12). Second, Lynch (2000a) has identified a "Southern Oceanic" subgroup (or rather a "linkage"), encompassing the languages of North-Central Vanuatu, Southern Vanuatu, and New Caledonia.[5] The family tree given in figure 1—a hybrid of various subgrouping studies since 1989—shows the relationship between the Southern Oceanic linkage referred to in this paper, and the wider "subgroup" of Eastern Oceanic as used by Geraghty (1990). Imperfect though it may be, this tree provides a convenient point of reference for the present study.[6]

Geraghty's study of *R identified 193 etyma containing *R for Proto-Eastern Oceanic (PEOc). Even more importantly, he also observed an important geographical pattern: "As a general rule, PEOc *R is lost in proportion to distance from Western Oceanic, beginning in the South-East Solomons" (Geraghty 1990: 90). The loss of *R, in Geraghty's account, took place in a few words in the Southeast Solomonic subgroup, in many more words in NCV, and in even more as one goes further south and/or further east in the Pacific. The Micronesian subgroup also lost *R in a number of words, and Central Pacific (Fijian, Rotuman, Polynesian) has lost almost all instances of *R.[7] In this article, I will be little concerned with Micronesian or Central Pacific languages, and shall restrict my observations to Vanuatu and New Caledonian languages.[8]

3. Although Geraghty speaks of "northern Vanuatu," he gives to this term the same meaning as what Clark (1985, 2009) calls "North-Central Vanuatu." For clarity's sake, I will therefore harmonize my terminology using the latter term, which is now widely used.
4. Geraghty also mentioned Southern Vanuatu languages (see 3.3) even though he did not formally include Southern Vanuatu and New Caledonia in his Eastern Oceanic subgroup.
5. Lynch's Southern Oceanic is distinct from the subgroup of the same name that Geraghty (1989) proposed for the languages of New Caledonia (Grande Terre + Loyalty Is). The latter subgroup is simply called New Caledonian by Lynch (2000a:158), after Ozanne-Rivierre (1992).
6. Following Ross (1988:41), a double line indicates (innovation-linked) "linkages" as opposed to (innovation-defined) subgroups.
7. The loss of *R is regularly cited as a shared innovation defining Proto-Central Pacific (Pawley 1972, 2007b:24; Geraghty 1983), even though Geraghty (1990:90) nuanced this general statement with respect to Fijian.
8. See, however, the short discussion in 4.4. In addition, appendix 3 will present firsthand data from the three languages of Vanikoro (Temotu, Solomon Islands).

FIGURE 1. A POSSIBLE FAMILY TREE FOR OCEANIC, SHOWING THE RELATION BETWEEN EASTERN OCEANIC AND SOUTHERN OCEANIC



Crucially, Geraghty showed that the geographic patterning is not only observable between each subgroup as a whole, but even *within* certain subgroups. Thus, for the North-Central Vanuatu subgroup, Geraghty (1990:85) categorized lexical items containing an etymological *R, according to their patterns of attested distribution across different areas of Vanuatu. According to him, "four isogloss bundles in Northern Vanuatu will account for most of the cases of loss of *R":

1. PEOc *R lost throughout North-Central Vanuatu [14 etyma],
2. PEOc *R lost between Mota and Raga [13 etyma],
3. PEOc *R lost between Paama and Namakura [4 etyma],
4. PEOc *R retained throughout North-Central Vanuatu [16 etyma].

The first three of these sets of isoglosses are represented in map 1. Set 2 is here arbitrarily represented as a line running south of the Banks Islands, even though Geraghty is less precise about its location: knowing that Raga is spoken in north Pentecost, the isogloss could actually sit anywhere between Mota and Pentecost.

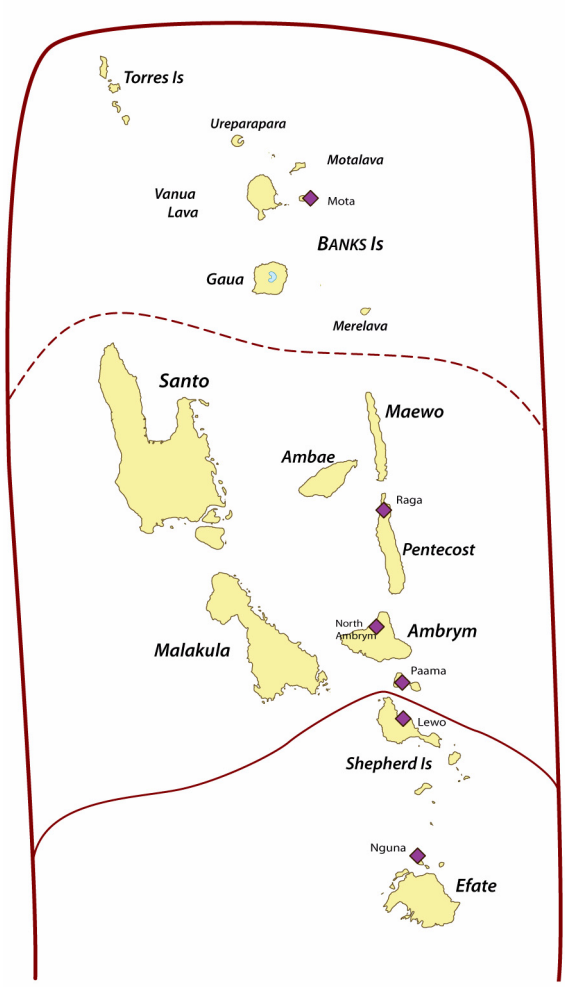
When Geraghty wrote his article, he had access to only a limited amount of published data: the major languages on which his study was based are indicated on map 1. For example, the northern area defined by the Torres and Banks islands, which will be the primary focus of the present study, was only represented in his sample by a single language, Mota—the only one that had then been properly documented (Codrington and Palmer 1896). For the same reason, Clark (2009), in his detailed reconstruction of the North-Central Vanuatu lexicon, chose Mota as the sole witness for the whole Torres-Banks area.

As he presented these rather neat isoglosses, Geraghty was aware that his representation of the facts might have been distorted by the gaps in the data: "It is possible, however, that

the siting of these isogloss bundles is an accidental consequence of the distribution of the better-documented and relatively conservative languages: Mota, Raga, [North] Ambrym, and Nguna" (Geraghty 1990:85). This cautionary footnote raises one question. What would the patterns look like if we brought data from previously undescribed languages from the same areas, so as to achieve a more fine-grained description of the loss of *R?

One possibility would be that the isoglosses proposed by Geraghty would be confirmed in a consistent way, in the sense that all languages within one area would behave like those that he chose in his sample; if so, such a tidy distribution could help, for example, in identifying neat subgroups within the NCV family. Alternatively, the introduction of

MAP 1. THREE ISOGLOSSES ILLUSTRATING LOSS OF *R IN NORTH-CENTRAL VANUATU LANGUAGES (after Geraghty 1990:85)



more fine-grained data might yield a much more random distribution of *R across languages, with retention only in some "conservative" languages, and loss in many others, with little geographical or dialectological consistency. Such a potential scenario would somehow disprove, at least partially, the conclusions proposed by Geraghty: his neat observations would then appear as mostly an artifact of his limited sample of languages.

New data can be found in Clark (2009), whose comparative lexicon takes into account more languages, especially from Santo, Malakula, and other islands further south. However, Clark's discussion of *R (2009:17) is brief, and does not mention any findings that significantly differ from those of Geraghty. In the present paper, I shall mainly refer to Clark's book for his reconstructions of Proto–North-Central Vanuatu (PNCV), and for the data—whether firsthand or secondhand—that are cited there.

Lynch (2009a) is a more detailed study of the loss of *R (and of *q), focusing on 11 Malakula languages. The data he brings are essentially consistent with Geraghty's earlier observations for the same etyma. He observes that most lexical items behave consistently across his sample, either losing or retaining *R in all 11 languages. Only a handful of lexemes seem to have a patchy distribution across Malakula, with no apparent regularity, whether phonetic or geographic: PNCV *m^{w}aRaki 'ground dove', POc *paRage '*Pangium edule*', POc *takuRu- 'back', and POc *suRuq 'juice'. With the exception of these very few items, Clark and Lynch neither contradict Geraghty's three-way partition of the NCV group, nor identify significant new geographical patterns within these. They thus leave the questions above open to further study.

1.3 AN INTERESTINGLY UNPREDICTABLE CHANGE. Although one can identify a few phonological factors—for example, the nature of the surrounding vowels—that may, as a tendency, have influenced the loss of *R, none seems to be absolute. As Lynch (2009a:62) put it, "it is not possible to define phonologically with any exactitude the conditions under which *R is retained or lost."

This irregularity will be confirmed in this study, especially through the observation of etymological doublets: that is, forms in which *R is both retained and lost, for the same language, in exactly the same environment (see 2.2.3). Therefore, apart from a short discussion in 3.4, the present article will not focus primarily on the phonetic conditioning for the loss of *R. Instead, the geographic distribution of *R retention and loss must be observed on a case-by-case basis—that is, considering each *R etymon separately.

The loss of *R thus appears to be irregular, in the sense that it does not obey any regular conditioning. Furthermore, if it was indeed a liquid of some sort, then the phonetic motivation for its deletion is not obvious: while it is cross-linguistically common for a liquid to change into another segment (for example, trill [r] > flap [ɾ], or > [l], or > [d], and so on), it is less common for it to disappear altogether—at least in comparison with notoriously unstable consonants like [h] and [ʔ].[9] Crucially, the fact that the loss of *R is both irregular and apparently little motivated[10] will make the geographical patterns of its loss all the more significant to the historian. Here is how Pawley (1972:30) comments on the "loss of *R" in various Oceanic languages: "The value of this loss for subgroup-

9. This being said, the deletion or weakening of *r is attested in various forms throughout the world, including in the Pacific, albeit rarely. For example, Proto-Polynesian *r > Ø in the Tongic languages, and > /ʔ/ in Marquesan (Marck 2000:23, Charpentier and François forthcoming).

ing purposes is somewhat uncertain because the behaviour of *R is less regular than that of most other consonants … Still, the loss of *R is relatively uncommon in Oceanic languages, and clearly carries much more weight for subgrouping than, say, the loss of *q."

When a sound change shared by several languages is phonetically easily explained (like *s > h, *^{m}b > p, *q > Ø, and the like), it constitutes weak evidence for any historical analysis—whether on subgrouping or diffusion—because parallel innovation always remains a possibility. But the situation with the loss of *R is different: because it doesn't seem to obey any regular phonetic conditioning nor any obvious motivation, and because it only affects parts of the lexicon for each language, it constitutes a potentially significant window onto the linguistic history of this part of Melanesia.

2. RETENTION VS. LOSS OF *R IN NCV LANGUAGES. In this data-oriented section, I aim at testing Geraghty's hypotheses about *R-loss in Eastern Oceanic. While I will first reduce the geographical scope to the North-Central Vanuatu area, I will also refine the grain of observation, by taking into account a denser set of linguistic data that are now available.

The principal outcome of these fine-grained observations will be to confirm the large-scale north-to-south cline observed by Geraghty. However, instead of defining three major zones within North-Central Vanuatu (as in map 1), the isoglosses related to the loss of *R will differ lexeme by lexeme, with at least fifteen different isoglosses for North-Central Vanuatu languages. Further sections (3.3) will later widen the scope to Southern Vanuatu and New Caledonia, showing that the same pattern can be extended throughout the Southern Oceanic linkage. In the final discussion (sections 4 and 5), I will examine how these empirical results shed new light on the history of language development in this part of Melanesia, and on the distribution of social networks at early times of settlement.

2.1 A PRELIMINARY NOTE ON SOURCES AND DATA. Before I begin to present the facts, I make the following preliminary note regarding the sources I use for the linguistic data, and the way they will be presented in the present article.

Throughout this paper, data will take the form of bundles of cognate forms, grouped under a given reconstruction, as in (1) below. I will discuss 92 different cognate sets, all of which involve an already established etymon containing *R according to existing sources. Apart from a couple of minor notes, I will not propose any new *R etyma. The source for PEOc etyma is Geraghty (1990). POc forms are taken from various sources, especially Ross, Pawley, and Osmond (1998, 2003, 2008, forthcoming). PNCV reconstructions are taken from Clark (2009).[11]

10. These are two different things. A sound change can be motivated by common principles such as lenition (like *p > *f or *f > *h), yet be irregular, that is, affect only some lexical items with no conditioning.

11. I normalize Clark's idiosyncratic orthography by transcribing *ŋ for his ⟨g⟩ and *g for his ⟨q⟩, in line with established usage for other protolanguages. Throughout this article, it will be implicit that all voiced stops (b, d, j, g) in protoforms must be reconstructed as prenasalized stops; thus *Rabia 'starch' stands for *Ra^{m}bia, *guRio 'dolphin' stands for *^{ŋ}guRio, and so on. But while I will follow traditional usage in leaving prenasalization implicit for reconstructions, I will make it explicit in the phonetic transcription of firsthand data (especially languages from the Torres and Banks islands), for the sake of phonetic accuracy.

The 92 cognate sets discussed in this paper are essentially comprehensive of all *R etyma attested among NCV languages. Several POc/PEOc reconstructions containing *R will not be considered in this section, because they are lost altogether from this area of reference (or at least from the Banks and Torres islands): *jiRi '*Cordyline*', *maRi 'breadfruit', *mawiRi 'left hand', *Ruqa 'neck', *paqoRu 'new', *qapaRa 'shoulder', and *waRoc 'creeper'. Some of these will be cited, albeit not systematically, in later sections relating to areas further south (3.3).

Within the North-Central Vanuatu linkage, the modern languages I will cite in this study are represented on map 2. The relevant data presented here will be of two kinds: firsthand data, collected by the author during various field trips to the Torres and Banks islands, between 1998 and 2007;[12] and secondhand data for languages further south, based on various publications, particularly Clark (2009), complemented with additional sources. Already published data will only be reproduced in this paper when appropriate; otherwise, reference will be given to the sources.[13] As for the firsthand data from the Banks and Torres islands, apart from a few cognate sets in François (2005), they are mostly unpublished, and therefore need to be presented comprehensively in this paper. When the modern reflexes of a particular etymon are relevant to the discussion, they will be indicated in the body of the paper; the rest of the supporting data from the Banks and Torres islands will be listed in appendix 1.

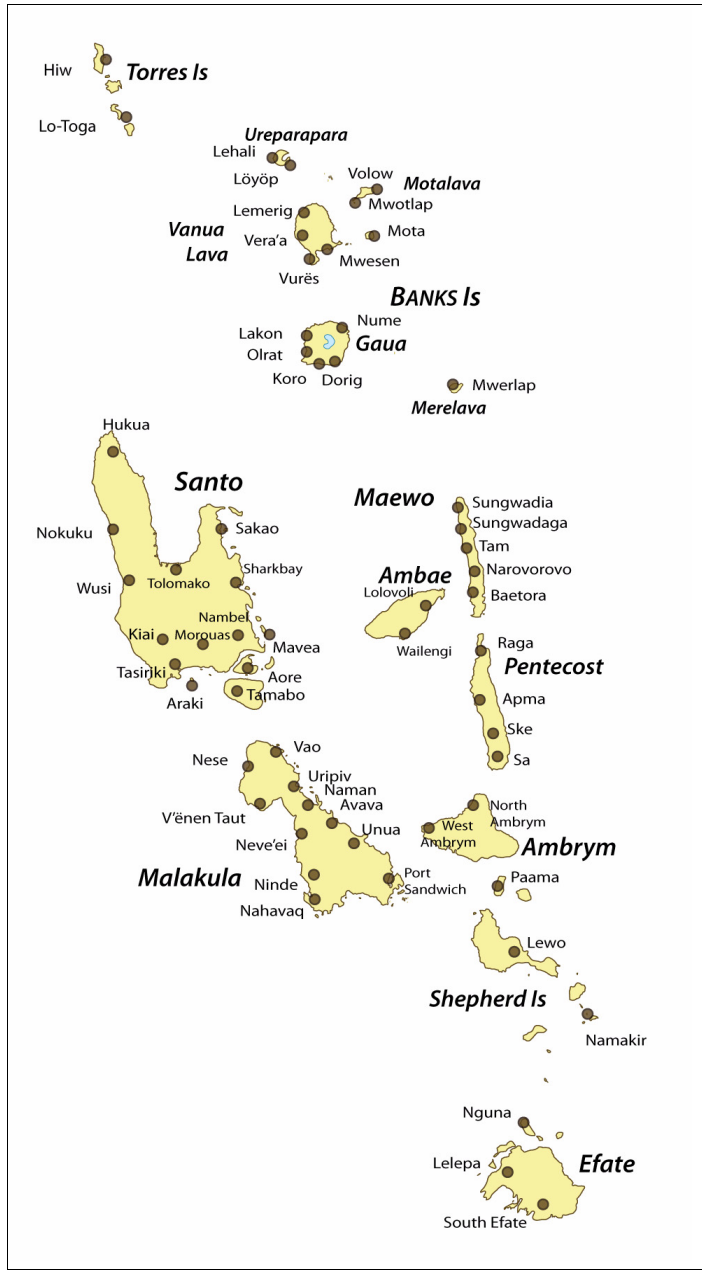
The Torres and Banks islands are home to 17 distinct languages (François 2005, forthcoming), the northernmost languages of map 2. In the data sets provided here, these languages will be listed following a consistent order running from northwest (Hiw) to southeast (Mwerlap); each language will be named using a three-letter abbreviation presented in appendix 4. All modern forms are given in IPA, except /ü/ for IPA [y]. Example (25) below, reproduced here as (1), is an example of such a cognate set.[14]

(1) POc *quRis > PNCV *uRis '*Spondias cythera*':
HIW ʉᵍʟ; LTG ʉr; LHI n-nuj; LYP n-øj; MTP n-ɪj; VRA n\ur; VRS ür; MSN ʊr; MTA$_1$ ur, MTA$_2$ us; NUM w\ɪs; DRG wa-ʊr; LKN u:; MRL nɛ-wɛ͡as. [Sungwadia is/a; Raga uhi; Nokuku us; Uripiv na-us.…]

As a rule, I shall reproduce here all the data I have from the Torres-Banks area. When a language is not cited for a given lexical item—like Lemerig or Olrat in (1)—this indicates either a gap in the data, or the lack of any reflex in the modern language due to lexical replacement.[15] These Torres and Banks data will sometimes be followed by a sample

12. I sometimes double-checked or complemented my own data with other sources, especially for flora and fauna terms: Codrington and Palmer (1896) for Mota, Hyslop (n.d.a) for Vera'a, Hyslop (n.d.b) for Vurës, and Schmidt (n.d.) for Nume.
13. These additional sources include Hyslop (n.d.c) for Lolovoli, Jauncey (n.d.) for Tamabo, François (2002) for Araki, Charpentier (1982) and Lynch (n.d.) for various Malakula languages, as well as personal communication from various linguists (see footnote 1).
14. Hyphens indicate morpheme boundary, for example, between article and noun. A slash indicates the limits of the reflex ('\' for left boundary, '/' for right boundary) when these are no longer morpheme boundaries in the modern language. Thus the Vera'a form *nur* is transcribed here as *n\ur*, in which the root coalesced with the former article (François 2007). Likewise, the Sungwadia form *is/a* demonstrably reflects *uRis with the accretion of a nonetymological (paragogic) *-a (see 2.3.3).
15. The contrast between MTA$_1$ and MTA$_2$ here refers to two dialects of Mota and, with respect to loss or retention of *R, is only relevant for this particular lexical item (see 2.3.3).

MAP 2. NORTH-CENTRAL VANUATU LANGUAGES CITED IN THIS STUDY



of languages further south: in this example, forms from Sungwadia, Raga, Nokuku, and Uripiv. Whenever these forms are representative of other NCV languages with respect to the loss or retention of *R, I will cite only two or three of them. Conversely, when the NCV situation is more complex—that is, when some languages lost *R while others retained it (2.3.3)—I will cite a larger number of relevant forms.

2.2 GENERAL PRINCIPLES ON THE FATE OF *R IN THE NCV LANGUAGES. Before we start to observe the geographical distribution of *R in Vanuatu, it is useful to discuss three general principles regarding the fate of this phoneme. First, 2.2.1 will examine the way in which *R, when it was not lost, is regularly reflected in NCV languages. Then 2.2.2 will discuss the phonotactic conditioning of *R-loss word-finally, and also exceptions to this principle. Finally, the question of etymological doublets will be the subject of 2.2.3.

2.2.1 Regular reflexes of *R. For the purpose of this paper, the most important point in the cognate sets will be to observe which languages reflect the retention of *R, and which ones show evidence that it was lost. For example, in (1) above, LTG *ur* clearly shows retention of the rhotic, whereas Uripiv *na-us* reflects the loss of *R. However, not all forms are so easy to interpret. For example, do such forms as MTP *n-ɟ* or LKN *u:* illustrate loss or retention of *R in *quRis? In order to be able to track with precision the fate of this protophoneme, it is necessary to state how exactly *R was regularly reflected in modern Vanuatu languages (when it was not lost).

In all NCV languages, all nonzero reflexes of POC *R are identical with those of POC *r (Pawley 1972:30, Tryon 1976:51).[16] This means that *R, at some point in its history, underwent a merger with *r (see 4.6), except for the words where *R became zero. In most languages, the modern reflexes of *r/*R are straightforward, typically an alveolar trill [r]. In Araki (François 2002) and Avava (Crowley 2006), *r/*R are reflected as an alveolar flap [ɾ]. In Paamese and Lewo (Lynch 2008), they are reflected as an alveolar lateral [l]. However, some modern languages show less obvious reflexes, which I propose to review here.

The following examples show that the Torres and Banks languages also show identical reflexes for *r and *R. Regular reflexes of *r are illustrated by (2), as well as by the forms in section 1 of appendix 1:

(2) POC *koro 'prepositional verb denoting motion around or against' > PNCV *koro 'surround, cover, obstruct; prevent; protect':
HIW **ᵍʟ**oɣ (met.); LTG **ɣo**r; LHI **ɣɛj**; LYP **ɣɔj**; VLW **ɣɔj**; MTP **ɣɔj**; LMG **ɣœ**r; VRA **ɣʊ**r; VRS **ɣʊ**r; MSN **ɣɔ**r; MTA **ɣo**ro; NUM **ɣɔ**r; DRG **ɣɔ**r; KRO **ɣɔ**r; OLR **wɔj**; LKN tu\wɔ:; MRL **ɣɔ**r.

Regular nonzero reflexes of *R are shown in (3)—repeating (33)—and will be widely confirmed by the other data cited in this study.

(3) POC *paRi > PNCV *vaRi 'stingray, *Dasyatidae* spp.':
HIW **βɔᵍʟ**; LTG **βɛr**; LHI n-**βæj**; LYP n-**βɪj**; VLW n-**βɛj**; MTP nɛ-**βɛj**; LMG n-**βɛr**; VRA **βɛr**; VRS **βœr**; MSN **βɛr**; MTA **βar**; NUM **βɛr**; DRG **βa:r**; KRO **β͡ɛar**; OLR **βaj**; LKN **βæ:**; MRL nɛ-**βɛr**.

16. In appendix 2, I mention only two aberrant cases where *R may be reflected as *l.

In 10 out of 17 languages (Lo-Toga, Lemerig, Vera'a, Vurës, Mwesen, Mota, Nume, Dorig, Koro, and Mwerlap), *r/*R is regularly reflected as an apical trill [r], in all positions.

In Lehali, Löyöp, Mwotlap, and Volow, *r/*R is systematically reflected as a palatal glide [j] in all positions. Incidentally, this regular change can be dated around the end of the nineteenth century (François 2001:62, forthcoming).

In Hiw, *r/*R has become a prestopped velar lateral [$\widehat{gL}$] in all positions (François 2010). I will henceforth transcribe this phoneme as /ᵍL/: POC *raqup > *rau > HIW [$\widehat{gL}$ɔ] transcribed *ᵍLɔ*. Hiw has complex rules of interference depending on stress (François 2010:419),[17] either when two occurrences of /ᵍL/ (< *r/*R), or when /ᵍL/ (< *r/*R) and /ɣ/ (< *k), occurred in adjacent syllables in the etymon. I will only cite three of these rules here (an underscore here represents a vowel).[18]

- metathesis: *'k_r_ > *'ɣ_ᵍL_ > 'ᵍL_ɣ_
  e.g., (35) *kiRe > *ɣire > *ɣiᵍLə > *ᵍLiɣə* 'pandanus'
- dissimilation: *'r_r_ > *'ᵍL_ᵍL_ > 'ᵍL_ɣ_
  e.g., (#5) *rarap > *ᵍLaᵍLə > *ᵍLaɣə* '*Erythrina*'
- assimilation: *k_'r_ > *ɣ_'ᵍL_ > ᵍL_'ᵍL_
  e.g., (41) *kaRuve > *ɣaruwe > *ɣuᵍLɯwə > *ᵍLɯᵍLɯwə* 'k.o. crab'

In a Hiw form like (35) *ᵍLiɣə* < *kiRe 'pandanus', the segmental reflex of *R is thus not /ɣ/ but /ᵍL/, with regular metathesis.

Regular processes of unstressed vowel reduction have triggered, in most languages of the area (François 2005), the reduction of disyllabic feet *CVCV to closed monosyllables C_1VC_2. In two languages of western Gaua, Olrat and Lakon, the protophonemes *r/*R show different reflexes depending on the rhotic's position in the newly created C_1VC_2 syllable.

In both languages, *r/*R > [r] syllable-initially, as in (#6) POC *rua 'two' > OLR, LKN ***r***ʊ, or (#42) POC *Ropok 'dash, fly' > OLR, LKN ***r***ɔw. But when *r ended up in a syllable coda, it is reflected as follows:

- In Olrat, syllable-finally *r > [j]. Thus (2) *koro > *ɣɔr > OLR *wɔ**j**;* (3) *paRi > *βar > OLR *βa**j**.*
- In Lakon, *r in syllable coda regularly disappeared with compensatory lengthening (François 2005:461), thereby triggering the phonemicization of vowel length in this language. Thus (3) *paRi > *βær > LKN *βæː* 'stingray' (contrasting with *βæ* 'thing').

A word-final long vowel in Lakon always points to the presence of a rhotic, whether *r or *R. This will be an important point when assessing the retention and loss of *R. For example, the long vowel */uː/* in (1) above illustrates a case of *retention* of *R (<*quRis). Similarly, consider the following minimal pair:

(4) a. LKN *$\widehat{kpʷ}$ɔ* with short [ɔ] 'pig' < *bʷóe, ultimately from (#9) POC *boRok, showing loss of *R.

b. LKN *$\widehat{kpʷ}$ɔː* with long [ɔː] 'dream' < *$\widehat{kp}$ʷɔr < *bʷóre, from (#26) PEOC *boRe, showing retention of *R.

17. I use "interference" in the sense of Blust (2009:206), to designate those cases where "segments are sensitive to one another in adjacent syllables."

18. Throughout this paper, a simple number in brackets, like (21), refers to the cognate sets that are given in the text. A number preceded by a sign '#', as in (#21), refers to the cognate sets in the appendices.

The various reflexes of *r/*R in NCV are summarized in table 1.

2.2.2 The rare retention of final *R. As mentioned in 1.3, no phonetic factor can be absolutely identified as a conditioning environment to explain the loss of *R in north Vanuatu languages—with one exception.

The only obvious regularity is that it was always lost in word-final position, as in words like *roŋoR 'hear', *pusuR 'bow', *qauR 'bamboo', and so on. Yet this behavior is not specific to *R, because etymological root-final consonants were systematically lost in North-Central Vanuatu languages (Clark 2009:17) whenever they were the last segment of a phonological word. Lynch makes similar observations with respect to the loss of root-final *R for Malakula languages (NCV) (Lynch 2009a:61) and for Southern Vanuatu (non-NCV) (Lynch 2001:40).

However, the regular deletion of word-final consonants can be shown to have taken place at a relatively late stage in the history of the NCV linkage (François forthcoming). Individual members of the dialect network (ancestors of modern languages) sometimes had developed suffixal morphology on some lexical items, thereby retaining the former root-final consonant before it was regularly lost. Thus a root-final *R was sometimes exceptionally retained, when it was followed by extra material within the same word. This is the case, for example, with the POC transitivizing suffix *-i, or the applicative *-aki[n]:

(5) POC *sinaR 'shine' + *-i > *siŋa**R**-i > MTA siŋa**r** 'illuminate' (see [#46])
POC *sinaR 'shine' + *-akin > *siŋa**R**-aki > *siŋaraɣi > MTP hiŋ**j**ɛɣ 'illuminate' (see [#46])

This case of retention is, however, sporadic. The root *roŋoR 'hear', though a transitive verb, has lost its final consonant *R everywhere; its combination with the applicative *-aki[n] shows a nonetymological epenthetic consonant /-t-/ > PTB[19] *roŋo-**t**aɣi (François 2005:482), instead of the expected rhotic. While (5) is unambiguously a case of *R-retention, it is unclear whether the absence of *R in a form like *roŋotaɣi should be analysed in the same way as other cases of *R-loss in the region, the focus of this study. Indeed, in Oceanic languages, nonetymological consonants at a morpheme boundary are not rare (van den Berg 2006) and sometimes result from later processes of morphological reanalysis. For this methodological reason, the present study will deliberately ignore cases like this one (root-final *R replaced by a different consonant at morpheme boundary) in the

TABLE 1. NONZERO REFLEXES OF THE PROTO-RHOTICS *r AND *R IN NORTH-CENTRAL VANUATU LANGUAGES

IPA	Reflex	Languages
r	alveolar trill	most of the 94 NCV languages
ɾ	alveolar flap	Araki (Santo), Avava (Malakula)
l	alveolar lateral	Paamese, Lewo (Central Vanuatu)
ᵍʟ	prestopped velar lateral	Hiw (Torres Is.)
j	palatal glide	four languages in north Banks + Olrat (Gaua) σ-finally
V:	lengthening of V	Lakon (Gaua) σ-finally

19. PTB stands for Proto–Torres-Banks: see appendix 1.

tracking of *R-loss. Note that their inclusion would not have affected the results anyway, due to their very small number—one or two examples only.

The extra morpheme that helped preserve *R can also be the final element in the circumfix *paRi-…-i 'unified action by a plural subject' (Pawley 1973:151):[20]

(6) POC *maturu(R) 'sleep' > HIW mitiᵍʟ 'sleep:SING.SUBJECT'
*(paRi-)matur**R**-i > HIW mɔtᵍʟiɣ 'sleep:PLUR.SUBJECT'

The Hiw reflex /ɣ/ here results from regular dissimilation of */mɔtᵍʟiᵍʟ/, as was discussed in 2.2.1.

In other cases, *R was supported by a nonetymological paragogic vowel *-i, which is attested in other Oceanic languages (for example, Clark 1985:204, Lynch 2000b:73):[21]

(7) POC *wakaR 'root' > *kawa**R**-i > LMG n-ɣœ**r** (see [#54])

In each case cited here, the extra vowel *-i may be absent from modern Torres-Banks forms, yet it must be posited in order to explain the quality of the preceding vowel (François 2005:479), as well as the unexpected retention of a word-final *R.

In all other cases, namely when *R was word-initial or word-medial in the etymon, the retention vs. loss of the rhotic obeys no obvious phonetic conditioning. This irregularity is especially evident from the observation of etymological doublets.

2.2.3 The case of etymological doublets. Several etyma containing *R are reflected in modern languages by two different words, with related yet distinct meanings. Crucially for this paper, one of the two words may show retention of *R, while its counterpart shows loss, and both forms may appear in the same language. These doublets confirm that the loss of *R cannot be straightforwardly assigned to any phonological conditioning (see 1.3), since the very same etyma have led to two different treatments of the consonant, despite the identical phonetic context.

For example, Mwotlap has two different reflexes of POC *meRaq 'red' (here combined with noun article): *nɛ-mɛ**j*** 'reddish color in the dawn sky', and *nɛ-mɛ* '*Bixa orellana*, plant from the seeds of which a red dye is extracted'. One might be tempted to analyze such examples as cases where the loss vs. retention of *R is "patchy." However, I believe this would not be a correct analysis. Despite their common etymology, these two forms are two distinct lexemes in Mwotlap. In terms of language comparison, they belong to two different cognate sets, whose semantic and phonological properties are independent of each other. On the one hand, the cognate set (8), repeated from (53), shows retention of *R in all known NCV languages:[22]

(8) POC *meRaq 'red' > PNCV *mera 'reddish color in the dawn or sunset sky':
LYP n-mɛ**j**; MTP nɛ-mɛ**j**; MTA me**r**a; … Tamabo me**r**a 'be a red sunset'

Conversely, the other set (9), also given as (#21), belongs to the category that shows consistent loss of *R (see 2.3.2):

20. Hiw sometimes retains the prefix *paRi- (see [24] below), yet with this verb 'sleep' it has disappeared. The same circumfix *paRi-… -i has left similar traces in the neighboring language Lo-Toga (François 2005:481).
21. The languages of Vanikoro have done the same in many lexemes, including some ending in *R: see appendix 3.
22. I here propose reconstructions at the level of PNCV in line with Clark (2009), even though the forms and/or the glosses are mine.

(9) POc *meRaq ‘red’ > PNCV *mea ‘red pigment; esp. *Bixa orellana*, a plant from the seeds of which a vivid red dye is extracted’:[23]
HIW mi̯e; LTG mi; MTP nɛ-mɛ; VRS mɪ; MTA mea ‘red earth’; DRG mɛ; LKN mɛ

The reason for the existence of such doublets is hard to identify. Late borrowing is a frequently cited explanation (Geraghty 1990:88, 2004), and cannot totally be ruled out here. If so, forms preserving *R would be later loans from conservative languages (?); or conversely, forms showing loss of *R could be borrowed from languages in which *R had been lost in that word. In both cases, one of the two forms would be a “direct” reflex of the etymon, while the other one would be an “indirect” reflex—that is, due to late borrowing. And indeed, hypotheses based on borrowing are sometimes convincing in the case of a cultural artifact or a plant, which are both the case here for *mea (see also *rava or *buraka in table 2). But it is less easy to figure out why a form for ‘red sunset’, for example, would be borrowed. Considering how the loss of *R is an irregular change anyway (see 1.3), the case for borrowing is not necessarily strong in all cases of doublets.

Now crucially, these etymological doublets should not be confused with those lexical items that show variation in the retention vs. loss of *R *within the same cognate set*. For example, the cognate set given in (1) above showed the *R of the etymon *quRis ‘*Spondias*’ to be retained in some languages, and lost in others. What (8) and (9) show is a different case, whereby each different lexical set shows internal consistency with regard to the loss and retention of *R.

This internal consistency makes it possible to tentatively reconstruct these etymological doublets for their common ancestor. Table 2 lists several cases of doublets or triplets at

TABLE 2. SOME POc ETYMA THAT HAVE SPLIT INTO ETYMOLOGICAL DOUBLETS IN PNCV

POc	meaning	PNCV	meaning	#
*meRaq	‘red’	*memea	‘red’	(#22)
		*mea	‘*Bixa orellana*, red dye’	(#21)
		*me**r**a	‘red color in dawn sky’	(53)
*tabiRa	‘bowl’	*tabia	‘wooden dish in which pudding is pounded’	(#18)
		*tabe**r**a	‘round, fine-woven basket used as a dish’	(#60)
		*tabe**l**a	‘triangular, coarse basket for rubbish’	(#60´)
*Rabia	‘starch’	*abia	‘starch, esp. sago starch’	(#20)
		***r**ava	‘Polynesian arrowroot’	(55)
*buRaka	‘swamp taro’	*buaga	‘taro swamp; boggy ground’	(#62)
		*bu**r**aka	‘swamp taro, *Cyrtosperma* sp.’	(#63)
*kaRat	‘bite’	*kati	‘chew (esp. kava)‘	(#19)
		*ka**R**a(ti)	‘bite; itch (lit., fig.), irritate, burn’	(23)
*tuRa-	‘sibling’	*tua-	‘companion, fellow’	(#14)
		*tuatua-	‘opposite-sex sibling’	(#15)
PEOc *ʔusuRi	‘follow; because of’	*(l)usu**r**i	‘[v.] follow, [prep.] along’	(#48)
		*su**r**i	‘because of (s.t.); because’	(#49)

23. The reconstructible sense is likely to be ‘red pigment’ in general, considering that *Bixa* sp. is a recently introduced species (Geraghty, pers. comm.).

the level of PNCV, using the conventions (*R > *r) set out in 2.3.2. The numbers in the right hand column refer to the data sets given later in this paper.

The existence of doublets is logically independent of the issue of loss vs. retention of *R. Thus table 2 shows some doublets that have both lost *R (< *tuRa-) or both retained it (< *ʔusuRi): even though they do not contrast with respect to the retention or loss of *R, they are still etymological doublets derived from an *R-etymon. Some cases, like *meRaq above, distribute reflexes of *R in two separate sets, each of which is internally consistent. Finally, *kaRat 'bite' combines the two dimensions: existence of a doublet, and internal variation in *R loss. Thus, one of the two lexical sets that reflect *kaRat has lost *R everywhere: POC *kaRat 'bite' > PNCV *kati 'chew (espec. kava)': HIW *ɣɔt*; LTG *ɣɛt*; LYP *ɣɪt*; MTA *ɣat* (see [#19]). Conversely, we will see that in the second set, (23), *kaRa(ti) shows both retention and loss of *R in North-Central Vanuatu (see 2.3.3). Among the various lexical sets cited in table 2, *kaRa(ti) is the only one that shows variation in *R-retention among NCV languages.

2.3 THE GEOGRAPHY OF *R-LOSS IN NCV LANGUAGES. I now turn to the central question of this study, namely the retention and loss of the protophoneme *R among North-Central Vanuatu languages, and its precise geographical distribution.

The present section will organize the data into three distinct classes. Although these are inspired by the categories used by Geraghty (see 1.2), they are distinct from them:

- 2.3.1 POC *R is lost throughout North-Central Vanuatu;
- 2.3.2 POC *R is retained throughout North-Central Vanuatu; and
- 2.3.3 North-Central Vanuatu languages show both loss and retention of POC *R.

The first two categories illustrate cases where *R was treated identically in all NCV languages; while these cases are relatively numerous, they are of minor interest to my research, and will be treated briefly. The last case will prove more interesting.

2.3.1 *R is lost everywhere in North-Central Vanuatu. I shall begin with those *R etyma that are reported by Geraghty (1990:85) and Clark (2009) to behave consistently across North-Central Vanuatu languages, by showing universal loss of *R.

Geraghty (1990:85) listed 14 lexical items in which *R appeared to have been lost throughout NCV. This was based on the data available to him at that time; as far as the Banks and Torres islands are concerned, this included only the language Mota. Examination of more data usually confirms Geraghty's treatment of these words. A first example is given in (10), repeating (#10):

(10) POC *piRaq > PEOC *viRa > PNCV *via '*Alocasia* taro':
HIW βi̯ə; LTG βiə; LHI n-βɛ; LYP n-βi͡ɛ; VLW nɪ-βɪ; MTP nɪ-βɪ; LMG n-βɪ; VRA βiɪ; VRS βɪ; MSN βɪ; MTA βia; NUM wɔ\βi; DRG βɪ; OLR βɪ; LKN βɪ; MRL nɛ-βe͡a. […Raga via; Araki via; Nguna na-via]

This situation can be formulated as a tendency (11):

(11) TENDENCY: When *R is lost in Mota *and* in languages further south, then *R is usually also lost in all other Torres and Banks languages.

In other words, Mota can be considered generally representative of its area. However, this tendency (11) shows some exceptions, which will be discussed in 2.3.3: these are words

that have lost *R in Mota and languages further south, yet have retained it in languages further north. A much more powerful prediction can be achieved if one starts not from Mota, but from the northernmost languages of the NCV group, namely the Torres languages:

(12) RULE: When a word lost *R in the Torres languages, then it lost it in all other NCV languages further south.

This observation (12) was illustrated by (10) above, and is confirmed by the 15 cognate sets cited in section 2 of appendix 1. There is only one exception (*vaRo '*Neisosperma oppositifolia*'), which will be discussed in 2.4.

2.3.2 *R is retained everywhere in North-Central Vanuatu. A second set of etyma are those where *R is consistently continued by a consonant throughout the North-Central Vanuatu linkage, and shows everywhere the same reflexes as the other rhotic *r.

Geraghty (1990:85) and Clark (2009:17) symbolize this pattern by reconstructing protoforms with *r at the level of PNCV, even though PEOc or POc had *R, for example, POc *(y)aRu > PNCV *yaru '*Casuarina equisetifolia*'. One way to read this is to consider that the *R of these words, at the time of the linguistic "unity" of PNCV, had already merged with the other rhotic *r. While my final interpretation will be slightly different (see 4.6), this formulation is fine at this stage of the reasoning. What precise sound change this shift is supposed to illustrate is, of course, difficult to determine, due to the unsolved question of the two rhotics' phonetic nature (see 1.1). But this formula *R > *r at least reflects the correspondence patterns observed within a given set of languages.

The observation of new data reveals that, when most documented NCV languages retained *R, then languages further north always did the same:

(13) RULE: When a word retained *R somewhere among NCV languages, then it retained it also everywhere else in the north, all the way to the Torres Is.

Apart from one exception already mentioned (see also 2.4), this rule (13) is overwhelmingly confirmed by the evidence available today. This is illustrated by the 37 different cognate sets given in section 3 of appendix 1. Among these, I here reproduce (#23) and (#26), respectively, as (14) and (15):

(14) POc *(y)aRu > PNCV *yaru '*Casuarina equisetifolia*':
HIW ɔᵍ**ʟ**; LTG ɔ**r**; LHI n-nɒ**j**; LYP n-ɪ**j**; VLW n-ɛ**j**; MTP n-ɛ**j**; LMG n-ɪ**r**; VRA nɛ**r**; VRS œ**r**; MSN ɔ**r**; MTA a**r**u. […Araki vi-aɾu; Uripiv n-ur; Namakir ne-ar…]

(15) POc *boRe > PNCV *bore 'dream':
HIW kʷoᵍ**ʟ**, kʷᵍ**ʟ**e; LTG kʷo**r**; LHI kʷɛ**j**kʷɛ**j**; LYP k͡pʷɔ**j**k͡pʷɔ**j**; VLW ᵑg͡bʷɔ**j**ᵑg͡bʷɔ**j**; MTP k͡pʷɔ**j**k͡pʷɔ**j**; LMG k͡pʷœ**r**; VRA k͡pʷʊ**r**; VRS k͡pʷʊ**r**k͡pʷʊ**r**; MTA k͡pʷo**r**e; NUM k͡pʷɔ**r**k͡pʷɔ**r**; DRG k͡pʷɔ**r**; LKN k͡pʷɔ:; MRL kʷœ**r**. […Araki popo**r**o; Uripiv bo**r**i; Namakir bo**r**…]

2.3.3 Etyma showing both retention and loss of *R in NCV. I now turn to those lexical items that vary in their retention of *R within the NCV linkage. This instability of the rhotic explains why Clark (2009) reconstructs it as *R at the level of PNCV: for example, (29) POc *taRaq > PNCV *taRa(ʔi) 'cut'. As we will see, this configuration corresponds to 32 different cognate sets.

Crucially, the data consistently delineate the same pattern, in the form of a north-to-south cline. This is summarized in (16).

(16) For each etymon containing *R, the loss of *R within the North-Central Vanuatu linkage is never random and patchy, but divides the archipelago into two neatly defined areas, one north and one south. The languages that retain *R are always located north of the *R isogloss, while those losing it are located south of it.

This observation (16) admits of two kinds of exceptions:

- sporadic loss of *R in an area where it was otherwise massively preserved (one word in 4 northern languages: see 2.4),
- sporadic retention of *R in an area where it was otherwise massively lost (a few words in a few languages of Malakula: see 4.3).

However, we shall see that these exceptions are few in number, and do not contradict in any significant way the massive tendency outlined here. Apart from these isolated cases, the principle in (16) holds true for all lexical items on which I have reliable data.

Interestingly, (16) confirms the observations that Geraghty (1990) had made on a larger scale, namely, the increased loss of *R as one goes further away from the Solomon Islands (see 1.2). However, because his survey at the time had a coarse grain, what seemed to emerge from his data were potentially neat isoglosses splitting the NCV archipelago into three sections, as in map 1. However, the second major observation that will appear from the data below is that the isoglosses defined by the loss of *R differ from lexeme to lexeme. The precise outlining of these different isoglosses is the object of this section.

In the cognate sets below, the boundary between the northern and southern isoglosses will be symbolized with a diamond ◆. All forms to the right of this diamond—including forms in NCV languages further south, which I will only partially reproduce here—show loss of *R. The isoglosses will eventually be represented on map 3.

2.3.3.1 Cases of *R-loss whose northern boundary runs within the Torres-Banks area. The name of the dolphin or porpoise contrasts the two Torres languages against the whole Banks group—and beyond this, against all NCV languages:[24]

(17) POC *kuRiap > PNCV *guRio ‘dolphin, porpoise’:
HIW k^{w}**ʁ**ʟɪ; LTG k^{w}u**r**iə; ◆ LHI n-kɛ ‘whale’; LYP n-kɪ ‘whale’; VLW nɪ-^{ŋ}gɪ; MTP nɪ-kɪ; VRS kɪ; MSN kɪ; MTA kio ‘whale’; NUM wi/ki; KRO kɪ; OLR kɪ; LKN kɪ; MRL nɛ-k͡ɛa. […Sungwadia kio ‘whale’; Raga ^{ŋ}gio; Paamese a-kio; Nguna giio…]

The area delineated by this contrast will be later represented on map 3 as isogloss number 2. I will here symbolize this as {*i2*}.

Geraghty (1990:85) counted the name of the *Canarium* almond, PEOC *ʔaŋaRi, among the words that had lost *R all across Vanuatu. This is because Mota, his main point of reference for the northern region, is located south of the actual isogloss. Thanks to more detailed knowledge of the languages of the area, we can now observe a neat divide between two groups of languages. The eight northernmost languages of Vanuatu

24. Those languages that have lost *R have reduced the resulting string *guio to a disyllable reconstructible as *gio.

retain *R in this etymon, whereas all languages further south have lost it. The isogloss boundary {*i3*} cuts right across the island of Vanua Lava, separating Vera'a from Vurës:

(18) POc *[ka]ŋaRi > PEOc/PNCV *ʔaŋaRi 'Canarium almond':
HIW ŋeᵍʟ; LTG ŋɛ**r**; LHI n-ŋæ**j**; LYP n-ŋɪ**j**; VLW n-ŋɛ**j**; MTP na-ŋɛ**j**; LMG n-ŋɛ**r**; VRA ŋa**r**; ◆ VRS ŋɛ; MSN ŋɛ; MTA ŋai; NUM ŋa; DRG ŋa; KRO ŋa; OLR ŋa; LKN aŋæ; MRL nɪ-ŋɪ. […Raga aŋai; Sakao aŋa; Namakir ʔaŋa…]

For this word, as for others, knowledge of regular phonological correspondences is necessary in order to observe such a neat divide. Earlier transcriptions of the Lehali form as /ŋæi/, or Mwotlap as /na-ŋei/ (Tryon 1976:290), might have led to the conclusion that these forms reflect the loss of *R, in much the same way as Mota /ŋai/; if this had been the case, the geographical distribution of the *R-loss would have appeared patchy and random. But research on the phonological history of these languages (François 2005) now makes it clear that the /i/ of the Mota form reflects the vowel of the etymon, which has here lost *R, whereas the /i/ of the Lehali and Mwotlap forms is in fact a palatal glide, the regular reflex of *R in these languages (see 2.2.1).

The cognate set for 'yesterday' defines a new isogloss {*i5*}. *R is retained in the eleven northernmost languages of the area down to Mota, but is lost from languages further south, both in southern Banks and in the rest of Vanuatu.

(19) POc *na-ñoRap > PNCV *nanoRa 'yesterday':
HIW nənɔᵍʟə; LTG nənɔ**r**ə; LHI nɒnnɒ**j**; LYP ɛnɔ**j**; VLW nɔnɔ**j**; MTP a\nɔ**j**; LMG lɪ\nɔ**r**; VRA nɔnɔ**r**ɔ; VRS nɔnɔ**r**; MSN lɛ\nɔ**r**; MTA a\nano**r**a; ◆ NUM nannɔ; DRG nannɔ; OLR nanɔnɔ; LKN nɔnɔ; MRL nananɔ. […Lolovoli nainoa; Araki nanovi; Nguna nanova…]

Note that some Banks languages, having lost *R in *nanoRa, subsequently gave the form more phonological substance by accreting another *na- syllable word-initially: thus *nanoa > *nananóa > NUM, DRG *nannɔ*; OLR *nanɔnɔ*; MRL *nananɔ*. In various NCV languages, a transitional /v/ appears between /o/ and /a/, that is, *nanoa > *nanova (Clark 2009:156); sometimes the final vowel changed to /i/.

The same isogloss {*i5*} is delineated by two other time adverbs, namely *waRisa 'the day after tomorrow' and *ana-waRisa 'the day before yesterday'. Obviously, (21) is derived from (20), at least historically, so it is possible that the loss of *R was linked in the two words.

(20) POc *waRisa > PNCV *waRisa 'the day after tomorrow':
HIW wuᵍʟi̯ə; LTG wu**r**iə; LHI **j**ɛh; LYP **j**ɛs; MTP ʊ**j**ɪh; VRS ɔ**r**ɪs; MSN wɪ**r**ɪs; MTA a**r**sa; ◆ NUM ais; DRG aŋɪs; LKN aɪh; MRL (mɛrɛntɪɣɪ). […Lolovoli waihe; Namakir pa-waih; Nguna waasa…]

(21) POc *qana-waRisa > PNCV *ana-waRisa 'the day before yesterday':
HIW nəwuᵍʟi̯ə; LTG nəwu**r**iə; LHI nɛnnɛ**j**ɛh; LYP ɛnɔ**j**ɛs; VLW nɪ**j**ɪh; MTP a\nɪ**j**ɪh; LMG lɪ\n**r**æs; VRS nɔnɔ**r**ɪs; MTA na**r**sa; ◆ NUM nais; DRG naŋɪs; LKN nahnɪ; MRL nɛwɛa͡s. […Lolovoli nawaihe; Nguna nanoasa…]

Those languages that lost *R as well as *w often ended up with a sequence of vowels: *anawaRisa > *(a)nawaísa > *(a)naaísa. In this case, the resulting vowel sequence was

sometimes broken by an epenthetic consonant: *-ŋ-* in Dorig (*naŋɪs* < *na**ŋ**aísa), *-n-* in Lakon (*nahnɪ* < *nanɪh* < *na**n**aísa). None of these consonants reflects *R.

The name of a beach creeper, '*Ipomoea pescaprae*' (POc *puRe) has been reconstructed by Clark as PNCV *vue-vue, based on the loss of *R in languages spoken south of the Banks. However, at least six languages in northern Vanuatu make it necessary to revise this PNCV reconstruction as *vuRe, because they retain *R:[25]

(22) POc *puRe > PNCV *vuRe 'beach creeper, *Ipomoea pes caprae*':
HIW ɵᵍʟə; LTG (hɔ) **ər**ə; MTP nʊ-jʊ\ʊ**j**; VRA ⁿdʊ-ɣa\βu**r**ʊ; VRS ⁿdʊ\ʊ**r**; MSN ɣa-nʊ\βʊ**r**; ◆ DRG ɣa\βʊβ. [… Nokuku wuwu; Paamese huahue…]

The northern limit of *R-loss here is located south of Vanua Lava—perhaps between Mota and languages further south, as in (19)–(21) above. Unfortunately, lack of data (including for Mota) makes it impossible to draw the line more precisely. However, the form *ɣaβoβ* in Dorig is likely to reflect *ɣa-βueβue, showing loss of *R just like in languages further south.

A fourth isogloss {*i6*} separates Mwerlap, as well as all languages further south (Clark 2009:109), from the sixteen Torres-Banks languages to their north:

(23) POc *kaRat 'bite' > PNCV *kaRa(ti) 'bite; itch (lit., fig.), irritate, burn':
HIW ᵍʟaɣəti (met.); LTG ɣa**r**əsi; LHI ɣa**j**; LYP ɣa**j**; VLW ɣa**j**; MTP ɣa**j**; LMG ɣɒ**r**-; VRA ɣa**r**-; VRS ɣa**r**; MSN ɣa**r**; MTA ɣa**r**a; NUM ɣa**r**ɛt; DRG ɣa**r**aːt; KRO ɣɛ**r**ɛat; OLR ɣa**r**at; LKN ɣæ**r**æs; ◆ MRL ɣɵt- ~ ɣat-. [Sungwadia kati; Hukua, Tamabo hati; Raga ɣasi; Unua xaʧi; Paamese ati; Nguna kati 'bite']

The next two examples are a bit more complex, as they seem to involve discontinuous areas, though in fact they still fit the pattern observed so far. First, the prefix *paRi- for reciprocal or plural action has preserved its *R in the twelve northernmost languages of the area (all the way down to Nume, in northeast Gaua). The next four languages on the list have all lost *R; yet the latter appears again, albeit optionally, in Mwerlap:[26]

(24) POc*paRi- > PNCV*vaRi- 'reciprocal action; plural action':
HIW βɔᵍʟ-; LTG β**ɛr**-; LHI β**æj**-; LYP β**ɪj**-; VLW β**ɛj**-; MTP β**ɪj**-; LMG β**ɛr**-; VRA β**ɛr**-; VRS β**ɛr**-; MSN β**ɛr**-; MTA β**ar**-; NUM β**ar**-; ◆ DRG βa-; KRO β-; OLR βa-; LKN βa-; ◆ MRL βa(**r**)-

What seems to be a patchy reflex becomes relatively tidy again if one realizes that Mwerlap is in fact geographically contiguous with Nume (see map 2). In other words, the *R-loss isogloss only encompasses the four southwestern languages of Gaua, as well as possibly the languages further south. Unfortunately, extant sources on Vanuatu languages do not seem to mention reflexes of *vaRi- in a way that would allow me to track the fate of *R in this prefix, or to draw a line on the map.[27] Geraghty (1990:85) includes it among

25. In several languages, the root *vuRe is preceded by the formative for 'leaves' (LTG *hɔ*, MTP *jʊ-*, VRA/VRS *ⁿdʊ-*, MSN *nʊ-* < PTB *ⁿrau- < POc *raun) or for 'vine' (VRA, MSN, DRG *ɣa-* < PTB *ɣa[w]e). Vera'a and Mwesen have both these prefixes, in different order (see François 2005:494–95 for these Proto–Torres-Banks reconstructions).

26. Some reciprocal forms in my Mwerlap corpus retain the *R, others do not. Information is missing on the conditions of alternation (*βa-* vs. *βar-*) in Mwerlap.

the items losing *R "between Mota and Raga," without providing the form. Clark (2009:70) cites it as *va(r)i- without providing the forms either, suggesting the rhotic is lost somewhere within the NCV group; I reconstruct this etymon as *vaRi-.

A similar situation holds for the name of the *Spondias* fruit, POC *quRis. *R is retained in all languages down to Mota, then lost in Nume, retained again in other Gaua languages, finally lost again in Mwerlap, and in all other NCV languages:[28]

(25) POC *quRis > PNCV *uRis '*Spondias cythera*':
HIW **ʉᵍɾ**; LTG **ʉr**; LHI n-nu**j**; LYP n-**øj**; MTP n-**ɪj**; VRA n**ur**; VRS **ür**; MSN **ʊr**; MTA$_1$ **ur**; ◆ MTA$_2$ us; NUM w\ɪs; ◆ DRG wa-**ʊr**; LKN u:; ◆ MRL nɛ-wɛa͡s. [Sungwadia is/a; Raga uh/i; Nokuku us; Uripiv na-us…]

Four Banks languages reflect an irregular form *oRi (instead of *uRi): these are LYP *n-øj*, MTP *n-ɪj*, MSN *ʊr*, and DRG *wa-ʊr*. These reflect irregular lexical changes that are common in the region (François forthcoming), and that are not directly linked to the loss of *R. More closely linked to the fate of *R is the observation that all Vanuatu languages that had lost *R in *uRis subsequently gave the form more phonological substance by accreting a nonetymological vowel (Clark 2009:17): thus *-i in Raga (*uRis > *uis-i > *usi > *uhi*) and most other languages, but *-a in Sungwadaga (*uRis > *uis-a > *wisa*), Sungwadia (*uis-a > *wisa > *isa*), and neighboring Mwerlap (*na uis-a > *na wisa > *nɛ-wɛa͡s*).

Interestingly for our dialectological survey, Codrington (1896) reports dialectal variation within the island of Mota itself: the Maligo dialect, located roughly north of Mota, has *ur*, but the Veverao dialect, located south, has *us*. In other terms, and somewhat surprisingly, the isogloss {*i4*} for *R-loss in this etymon *quRis cuts through the tiny island of Mota. From there it runs southeast and encompasses Nume and Mwerlap on its way to southern NCV languages, yet leaves the rest of Gaua untouched.

2.3.3.2 Cases of *R-loss whose northern boundary runs somewhere between the Banks Islands and Efate. The gradual north-to-south cline just illustrated for Torres and Banks languages continues further south. The following paragraphs will review those lexical items whose *R-loss boundary (symbolized by a diamond '◆') is situated south of the Banks islands. Contrary to previous pages, I will here provide more detailed data from languages outside the Torres-Banks area in order to help locate isoglosses precisely. I will mostly illustrate the languages closest to the isogloss boundary, that is, the southernmost language(s) having retained *R, and the northernmost language(s) having lost it. When no further indication is given, it must be understood that the rest of the data available are consistent with this presentation; that is, the languages cited after the diamond are representative of all NCV languages located further south.

The situation in the island of Malakula will be summarized in square brackets at the end of some cognate sets; these data come from Lynch (n.d.). A few languages of

27. The next language further south, namely Sungwadia (north of Maewo Island), has a reciprocal prefix /vaɣala-/ (Henri 2010:351), which does not seem to be cognate with *paRi-.
28. Clark (2009) proposes *uRi-si as his reconstruction of the PNCV form, instead of the expected form *ʔuRis. The absence of the expected glottal stop is an artifact of his methodology (2009:17), whereby /ʔ/ is only reconstructed when Namakir has kept a reflex. Although the proper reconstruction might warrant a debate, the presence or absence of the glottal has no bearing upon our discussion of *R. The quasi-universal loss of *q (*q > *ʔ > Ø) in NCV languages is an independent change anyway.

Malakula sometimes constitute exceptions to the neat north-to-south cline observed elsewhere. Once again, these exceptions differ lexeme by lexeme, and cannot be generalized (cf. Lynch 2009a). I will mention them in the cognate sets below, and come back to them in the later discussion (4.3).

Sets (26)–(29) are clear cases where *R-loss begins immediately south of Mwerlap, that is, between Mwerlap and Sungwadia (Henri 2010), the northernmost language of Maewo:

(26) POC *kuRita > PNCV *kuRita 'octopus, squid':
HIW ᵍʟitə; LTG ɣəritə; LHI n-wɛjɛt; LYP n-wujɛt; VLW nɪ-wɪjɪt; MTP na-wjɪt; LMG wiri?; VRA wiri?ɪ; VRS wɪrɪt; MSN wɪrɪt; MTA wirita; NUM wirit; DRG wrɪt; OLR wurɪt; LKN wɪrɪt; MRL nɪ-wɪr͡ɛat ◆ Sungwadia wita; Raga ɣuita; Hukua huita; Sakao neð; Araki huira; Nguna wiita; Lelepa wiit; [*R lost in Malakula]

(27) POC *draRaq > PNCV *daRa 'blood':
HIW taᵍʟə; LTG ʈarə, ʈərɛ-; LHI n-ⁿdaj; LYP n-ⁿdaj; VLW na-ⁿdaj; MTP na-ⁿdaj; LMG tɒr, tœrœ-; VRA ⁿdara; VRS ⁿdar; MSN nar; MTA nara; NUM ⁿdar; DRG ⁿdar; OLR ʧara, ʧari-; LKN ʧæræ, ʧarɪ-; MRL na-ⁿdar ◆ Sungwadia, Wailengi ⁿdai-;[29] Wusi ʧae; Tamabo ⁿdai; Araki ʧai; Nguna na-daa; [*R lost in Malakula]

(28) POC *takuRu 'back' > PNCV *takuRu 'back, behind; afterwards, later':
HIW tiᵍʟiɣ (met.); LTG təɣur; LHI tuɣuj; LYP ʧiɣij; MTP tiɣij; VRS təwür; MSN ɣɔr\tuwur; MTA taɣir ~ tawur; NUM a\tuwur; DRG twur ~ tawri; LKN tawu: … ◆ Sungwadia tawu- 'back'; Wusi tau-; Raga a-taɣu-; Nguna na-taku; [*R lost in Malakula, except possibly in two languages]

(29) POC *taRaq > PNCV *taRa(?i) 'cut (wood+), chop':
HIW taᵍʟə; LTG tarə; LHI taj; VLW taj; MTP taj; LMG ʔɒr; VRA ʔara; VRS tar; MSN tar; MTA tara; NUM tar; DRG tar; KRO tar; LKN tæræ; MRL tar ◆ Sungwadia, Raga, Tamabo tai; Araki rai; Uripiv e-tai; Namakir taʔ; Lelepa ta-i; [but *R is retained in six languages in Malakula]

For these four words, *R is preserved in the 17 Torres-Banks languages, and lost everywhere else in NCV: see the isogloss {*i7*} on the map.

For other lexical items, the isogloss runs further south. I first list those sets in which existing data make it possible to draw the isogloss with geographical precision. Thus (30) *suRi loses its *R in the middle of Maewo {*i8*}:

(30) POC *suRi > PNCV *suRi 'bone':
HIW siᵍʟ, sᵍʟi-; LTG hur, hʉri-; LHI n-huj; LYP n-sʉj, n-sijɨ-; VLW ni-hij; MTP ni-hij; LMG siri-; VRA siri-; VRS siri-; MSN siri-; MTA suri-; NUM sur; DRG sri; KRO siri; LKN suri, sʊrɪ-; MRL sur, suri-; Sungwadia suri-; … Tam suru- ◆ Narovorovo si-; Hukua, Nokuku, Araki sui-; Namakir siw; [*R lost in Malakula]

Example (31) shows loss of *R in *suRuq between east Ambae and north Pentecost {*i9*}:

29. In some languages, like Olrat and Lakon, the unsuffixed form *daRa alternates with a suffixed variant *daRi-. This nonetymological vowel *-i is also present in the form *ⁿdai-* recorded in several languages south of the Banks.

(31) POc *suRuq 'juice' > PNCV *suRu 'juice, fluid', *suRu(i) mata- 'tears': HIW tu^{g}**ʟ**-mətɔ-; LTG h**u**r-mətɔ-; LYP n-s**uj**-møʃø-; MTA su**r**-mata-; NUM su**r**-mata-; MRL su**r**-mata-; … Baetora su**r**i-mata-k 'tears'; Lolovoli su**r**u 'snot, mucus'; ◆ Raga hu- 'oil, liquid, juice, semen'; Sa su '(bodily fluid) run'; V'ënen Taut ui- 'fluid'; Uripiv suwe- 'juice'; Paama sii- 'juice' [NB: no data for Santo; possibly pockets of retention in Malakula]

Example (32) shows *Roʔoti(k) losing its *R between Raga and Abma, on the island of Pentecost {*i10*}:

(32) PEOc *Ro(ʔo)tik > PNCV *Roʔo(ti) 'bind, tie, bundle': HIW ^{g}**ʟ**øt; LTG **r**ət; LYP **j**øt; MTP **j**ɪt; VRS **r**øt; MTA **r**ot; MRL **r**ət; […] Raga **r**osi 'carry on stick'; ◆ Abma wootsi; Maëa, Tamabo oti; Araki ori 'a bundle'; Ninde us; Namakir ʔot 'tie a knot'; Nguna m\ooti 'bundle' [NB: no data from northern Santo; possibly pockets of retention in Malakula]

The line {*i11*} of *R-loss in (33) *paRi cuts again through the island of Pentecost, but this time further south—between Ske and Sa. Thanks to the data in Tryon (1976: 266), it is also possible to trace with some precision the line running across Santo:

(33) POc *paRi > PNCV *vaRi 'stingray, *Dasyatidae* spp.': HIW βɔ^{g}**ʟ**; LTG β**r**; LHI n-β**j**; LYP n-β**j**; VLW n-β**j**; MTP nɛ-β**j**; LMG n-β**r**; VRA β**r**; VRS β**r**; MSN β**r**; MTA β**r**; NUM β**r**; DRG βa:**r**; KRO β͡ea**r**; OLR β**j**; LKN βæ:; MRL nɛ-β**r**; … Tolomako β**r**i; Raga, Wailengi (East Ambae) β**r**i; Ske ko\fe**r**; ◆ Sa ea (?); Morouas a\fai; Araki a\vai; Maëa ɣo\ëai; Uripiv nö-vi; Paamese a-haai; Nguna vaai; [*R lost in Malakula]

Example (34) *Rapu(n) does essentially the same in Pentecost, but runs further south-east within Santo {*i12*}:

(34) POc *Rapu(n) 'haze, mist' > PNCV *ma-Ravu 'haze, mist' ~ *ma-RavuRavu 'hazy, blurred': HIW mə^{g}**ʟ**ɔw; LTG mə**r**əw**r**ɔw; LYP m**j**øp**j**øp; MTP mɪ**j**ɪp**j**ɛp; MTA ma**r**aβ**r**aβ; DRG m**r**aβ**r**ʊβ; LKN ma**r**aβ**r**aβ; … Raga ma**r**av; Wusi a^{n}**d**^{r}aβu; Tasiriki ʔa-**r**avu; Shark Bay ʔa-**r**avu; Morouas a^{n}**d**^{r}aβu; Nambel a^{n}**d**^{r}af; ◆ Maëa eu; Aore meu; Sa mau; Lewo mao; Nguna na-mavu; [*R lost in Malakula]

The line of (35) *kiRe {*i13*} runs somewhere between Malo (for Tamabo) and Northern Malakula (for V'ënen Taut):

(35) POc *kiRe > PNCV *kiRe '*Pandanus tectorius* (especially used for weaving mats)': HIW ^{g}**ʟ**iɣə (regular met.); LTG ɣi**r**ə; LHI n-ɣɛ**j**; LYP n-ɣ͡iɛ**j**; VLW nɪ-ɣɪ**j**; MTP nɪ-ɣɪ**j**; VRS wʊ\ɣɪ**r**; MSN ɣɪ**r**; MTA ɣ**r**e; MRL nɪ-ɣɪ**r**ɪt; Sungwadia xi**r**e; Raga ɣi**r**e; Tolomako hi**r**e 'pandanus mat'; Tamabo hi**r**ehi**r**e 'woven container' ◆ V'ënen Taut hei 'mat'; Nguna a-kie 'pandanus mat'; Lelepa na-ki

In (36), *vaRa(si) loses its *R between Ambrym and Paama. More precisely, the boundary {*i14*} cuts across the islands of Malakula and of Paama:

(36) PEOc *vaRa(cz) > PNCV *vaRa(si) 'tread on, step on':
HIW βaᵍʟə; LTG βar**ə**; VLW βa**j**; MTP βa**j**; LMG βɒ**r**; VRS βa**r**; MSN βa**r**; MTA βa**r**a ~ βa**r**as; LKN βæræh … Sungwadaga, Tamabo va**r**asi; Raga va**r**a; Abma va**r**ih; … West Ambrym ve**r**e; … Naman ve**r**es; Nahavaq βʲa**r**; South Paamese he**l**asi ◆ North Paamese heasi; Port Sandwich hee; Namakir baʔah

For some words, the lack of data for all the intermediate languages (here symbolized by '[…]') makes it difficult to draw a precise line on the map. Yet, crucially, even those cognate sets show a clear north vs. south contrast; and the distribution of *R-reflexes is often consistent with those words that can be tracked with more precision. For example, the following seven cognate sets (37)–(43) match closely the territory defined by (26)–(29) above, since they contrast the Torres-Banks area with islands further south—cf. {*i7*}.

(37) POc *Ruap > PNCV *Rua(vi) 'rising tide':
HIW (jɔβə); LTG (liɔβə); LHI **j**ɔ; LYP **j**ʊ; VLW **j**ʊ; MTP **j**ʊ; LMG **r**u; MTA **r**uaβ; NUM **r**uɛβ; DRG **r**ɪβ; MRL **r**uɛp; […] ◆ […] Raga (Geraghty 1990, no form given); Sa u; Wusi, Maẅea, Tamabo ua; [*R lost in Malakula, e.g., Nahavaq wu]

(38) POc *cakaRu > PNCV *sakaRu 'reef':
MTP na-skɛ**j**; VRS sœkœ**r**; MSN sɔkɔ**r**; MTA saka**r**; MRL nə-sɛk͡ɔə**r** […] ◆ […] Raga (Geraghty 1990, no form given);[30] Lolovoli sagau; Neve'ei ne-saʔau; Port Sandwich saxao; Nguna na-sakau; Lelepa na-skau; [no data for Santo; *R lost in Malakula, e.g., Nahavaq na-ʔhaw]

(39) POc *d[r]aRaka 'wild nutmeg, *Myristica fatua*':
MTP na-ⁿd**j**aɣ; VRA ⁿda**r**aɣa; VRS ⁿda**r**aɣ; MSN wɔ\na**r**aɣ; MTA na**r**aɣa; NUM ⁿda**r**ɛn; […] ◆ […] Sungwadaga dadae, Lolovoli dadai; Raga oaɣa; [no data for Santo; *R lost in Malakula, except Uripiv drrari]

(40) PNCV *daweRu 'coconut crab, *Birgus latro*':
LHI n-ⁿdɪ**j**; MTP na-ⁿdi**j**; LMG tɪ**r**; VRA ⁿdiɪ**r**; VRS ⁿdɪ**r**; MSN nɪ**r**; MTA nae**r**; DRG ŋ\ⁿdɪ**r**; KRO ⁿdɪ**r**; OLR ʧɪ**j**; LKN ʧɪː […] ◆ […] Maẅea daiu; Araki ʧau; Tamabo ⁿdaui; Lolovoli daweu; Raga davweu; [*R lost in Malakula, e.g., Nahavaq ⁿduwi, but Unua ruer]

(41) POc *kaRupe > PNCV *kaRuve 'ghost crab, *Ocypode*':
HIW ᵍʟʉᵍʟʉwə (regular assimil.); LTG ɣə**r**ʉwə; LHI n-ɣɔ**j**ɔw; LYP n-ɣa**j**ʊw; MTP na-ɣ**j**ʊw; VRS ɣø**r**üm; MTA ɣa**r**we; DRG ɣ**r**ʊw; MRL wʊ**r**ʊw; […] ◆ […] Sa awe; West Ambrym au; Unua, Port Sandwich xauv; Nguna kaapwe; [no data for Santo; *R lost in Malakula]

30. The official name of Sakao Island, northeast of Santo, reflects *sakaRu 'coral reef' with loss of *R. However, the island's local name is an unrelated form Laðhi; the form *Sakao* (or *Sakau*?) comes from another language, presumably one spoken in the vicinity (Raga? Tolomako?). The loss of *R illustrated by this name is therefore not to be assigned to the language Sakao, but to that other, unidentified language.

(42) POc *tapoRa 'nut-bearing tree' > PNCV *tavoRa '*Terminalia catappa*':
MTP na-tw**ɔj** '*Albizia saman*'; MTA tawo**r**a 'k.o. tree'; […] ◆ […] Sungwadaga, Maẅea, Raga, Lolovoli tavoa; Tangoa vi-tavoa; Abma towo; Namakir tawo; [*R lost in Malakula]

(43) PEOc *taRam-i > PNCV *taRam[an]i 'answer a call':
MTP ta**j**-ɣɔj; VRA ʔa**r**am 'tell s.o. s.th.'; MTA ta**r**ama; […] ◆ […] Sungwadia tami; Wusi, Nokuku tami 'allow, agree'; Raga dami 'agree'; Abma dam 'answer'; Navwien dam 'accept'; [*R lost in 4 Malakula languages, retained in 9]

Interestingly, the latter isogloss *{i7}* includes several forms that show the same phonetic profile *DaRa, in which "D" stands for a dental or alveolar consonant: *taRaq, *taRam(i), *draRaq, *d[r]aRaka. This may be one of the rare cases where the loss of *R may be regularly conditioned, for the same set of languages (but see 3.4).

The following two sets seem to outline the same area *{i8}* as (30) *suRi above:

(44) POc *paRage > PNCV *vaRage '*Pangium edule*; dance rattles':
HIW βa^{g}**ɹ**ak; LTG βə**r**ak; LYP n-βa**j**aŋ; MTP nɔ-wɔ\p**j**ak ~ na-p**j**ak; VRA wʊ\βa**r**ak; VRS wɪ\βi**r**i͡ak; MTA βa**r**ake; DRG wa\β**r**ak; LKN βæ**r**æk; MRL nɘ-βɛ**r**ɛ͡ak; Sungwadia fa**r**aki; Sungwadaga va**r**a^{ŋ}ge […] ◆ Lolovoli vake; Raga va^{ŋ}ge; Sa waak; [*R lost in Malakula: e.g., V'ënen Taut na\ẅak]

(45) POc *paRu > PNCV *vaRu '*Hibiscus tiliaceus*':
HIW βɔ^{g}**ɹ**; LTG βɔ**r**; LHI n-βɒ**j**; LYP n-βɪ**j**; VLW n-βɛ**j**; MTP nɛ-βɛ**j**; LMG n-βɛ**r**; VRA βɛ**r**; VRS βœ**r**; MSN βɔ**r**; MTA βa**r**; DRG βa:**r**; KRO βɛ͡a**r**; OLR ʧa-βa**j**; LKN βa:; MRL nɔ-β͡ɔɘ**r**; Sungwadia fa**r**u; Sungwadaga va**r**u […] ◆ […] Lolovoli, Nokuku vae; Kiai vai; Shark Bay θa; Araki vi-ẅa; [*R lost in Malakula]

Example (46) matches the area *{i10}* of (32) *Roʔo(ti) above:

(46) PEOc *(c,z)iRi(v) 'cut' > PNCV *siRi(vi) 'grate surface w. knife; peel; shave':
LTG hə**r**hi**r** 'shave'; LHI si**j**; LYP si**j**; VLW si**j**; MTP si**j** 'peel', sisi**j** 'shave'; VRS si**r** 'shave'; MSN si**r** 'grate vegetable'; MTA si**r** 'shave'; NUM, DRG si**r** 'grate coconut'; LKN hi: 'grate coconut'; Sungwadia si**r**i; Lolovoli hi**r**i […] ◆ […] Kiai zivi-a 'cut with knife'; Lelepa sii

Example (47) roughly matches the area *{i13}* of (35) *kiRe:

(47) PROc *b^{w}akaRe 'porcupine fish, *Diodon* sp.':
LTG k^{w}əɣa**r**; MTP na-k͡p^{w}ɣa**j**; LMG k͡p^{w}aɣa**r**; VRA k͡p^{w}aɣa**r**; VRS k͡p^{w}aɣa**r**; MSN k͡p^{w}aɣa**r**; MTA k͡p^{w}aɣa**r**e; DRG k͡p^{w}ɣa**r**; LKN k͡p^{w}aɣæ:; […] Nokuku poka**r** […] ◆ […] Nahavaq no-^{m}bowʔai; Namakir b^{w}aka; Lelepa k͡p^{w}okai

The set in (48) *kaRi[31] runs much further south *{i14}*, in a way similar to what we saw for (36) *vaRa(si):

31. Clark (2009:109) reconstructs this etymon as PNCV *ka(r)i, I have revised it to *kaRi.

(48) PEOc *kaRi > PNCV *kaRi ‘shellfish; esp. bivalve used as scraper’:
LHI n-βin\ɣæ**j**; MTP ni-βin\ɣɛ**j**; LMG ɣɛ**r**; VRS wɛ\ɣɛ**r**; MTA ɣa**r**; NUM wɔ\ɣɛ**r**; DRG ɣaː**r**; LKN ɣæː; MRL nɛ-ɣɛ**r**; Sungwadia xa**r**i; Wusi ʔa**r**i; Maẅea a**r**i; Tamabo ha**r**i; Raga ɣa**r**i; Uripiv n-a**r**; Vao na-ɣa**r** ‘k.o. shellfish’; […] ◆ […] Nguna kaai

For other words, attested forms are geographically too spread apart—either due to inadequate documentation or to lexical replacement—to allow for any placement of an isogloss boundary, even vaguely. However, once again, the general north-to-south orientation is confirmed.

(49) PEOc *kaboRa > PNCV *kaboRa ‘catfish’:
MTA ɣa$\widehat{\text{kp}}^{\text{w}}$o**r**a […] ◆ […] Uripiv n\abo; Lewo kapwo; Namakir ikapo

(50) POc *taRu ‘cover’> PNCV *taRu ‘cover; bake food in stone oven overnight’:
HIW tɔ$^{\text{g}}$**ʟ**; LTG tɔ**r**; LHI tɒ**j**; LYP tɪ**j**; VLW tɛ**j**; MTP tɛ**j**; VRA ʔɔ**r**; VRS tœ**r**; MSN tɔ**r**; MTA ta**r**; MRL tɘ**r**; […] ◆ […] Lelepa tau ‘bake in stone oven’

(51) POc *taRutu(m,ŋ) > PNCV *taRutu ‘porcupine fish, *Diodon* spp.’:
HIW tə$^{\text{g}}$**ʟ**it; VRS tɪ**r**ɪt; MTA te**r**it; DRG t**r**it; OLR ti**r**it, wʊ\tɪ**r**it**r**it; MRL nɛ-tɛ**r**it; […] ◆ […] Maẅea taut; Uripiv, Unua daut

Finally, some etyma retain *R everywhere in the northern languages, but my sources do not provide any other reflex in NCV languages further south—at least none showing loss of *R:

(52) PEOc *buRe > PNCV *b$^{\text{w}}$ure ‘ignorant, fool’:
HIW k$^{\text{w}}$u$^{\text{g}}$**ʟ**ə; LHI k$^{\text{w}}$ɔ**j**; LYP $\widehat{\text{kp}}^{\text{w}}$ʊ**j**; MTP nʊ-$\widehat{\text{kp}}^{\text{w}}$ʊ**j**; VRS $\widehat{\text{kp}}^{\text{w}}$ʊ**r** ‘forget’; MTA kp$^{\text{w}}$u**r**e

(53) POc *meRaq ‘red’> PNCV *mera ‘reddish color in the dawn or sunset sky’:
LYP n-mɛ**j**; MTP nɛ-mɛ**j**; MTA me**r**a; … Tamabo me**r**a ‘be a red sunset’. Cf. doublets *mea (#21) and *memea (#22).

(54) POc *naRa > PNCV *nanara ‘*Pterocarpus indicus*’:
HIW na$^{\text{g}}$**ʟ**ə; LTG ni$\widehat{\text{e}}$**r**ə; VLW na-na**j**; MTP na-na**j**; VRA nana**r**a; VRS nana**r**; MSN nana**r**; MTA nana**r**a; DRG nna**r**; LKN nana:; … Raga nana**r**a

(55) POc *Rabia ‘starch’ > PNCV *rava ‘Polynesian arrowroot, *Tacca leontopetaloides*’:
HIW $^{\text{g}}$**ʟ**aβə; LTG **r**aβə; MTP na-**j**ap; MSN **r**aɸ; MTA **r**aβa; LKN **r**aβ. Cf. doublet (?) *abia (#20).

It is ambiguous whether these cases illustrate the same north-to-south cline as above, or if they belong to those words that have kept *R everywhere (2.3.2). To these few examples, one may add a handful of sets (#60)–(#63) that will be discussed in appendix 2.

2.4 A SINGLE EXCEPTION. Whenever *R was retained (as *ɾ) in any given language of North-Central Vanuatu, the rule was always for all languages spoken *north*

of that point to show the same retention, as per (16) above. Overall, this north-to-south cline in *R-loss has been confirmed by 35 distinct cognate sets, that is, (17)–(51). The next section will discuss this phenomenon, which necessarily warrants some form of unified explanation.

This distribution suffers two kinds of exception. On the one hand, some languages of Malakula located south of the boundary show pockets of retention; I will come back to them in 4.3. On the other hand, one lexeme shows the reverse situation, namely a pocket of *R-loss in a region dominated by *R-retention: this is the name of the twin apple, PROC *vaRo.[32]

(56) PROC *vaRo > PNCV *vaRovaRo 'twin apple, *Neisosperma oppositifolium*':
HIW βa; LTG βəβa; MTP na-βaβa; ◆ VRA βa**r**βa**r**; VRS βa**r**βa**r**; MTA βa**r**oβa**r**o; DRG wβa**r**βa**r**; LKN βa:βa: ; Raga va**r**ova**r**o; Araki ð̪a**r**a ð̪a**r**a; Uripiv (bi)-va**r**va**r**; Nguna na-va**r**iva

Within North-Central Vanuatu, this is the only word that shows such a geographical pattern, whereby *R is retained in the south but lost in some languages further north. What is important here is not this pattern per se, but its exceptional rarity. It obviously reflects a local innovation, not linked with the set of *R-loss innovations that have spread from the south (see also 4.4).

3. A NORTH-TO-SOUTH CLINE. The previous sections consisted of the careful charting of the reflexes of Oceanic etyma containing *R in languages of north and central Vanuatu. The results of these observations can now be summarized.

3.1 SUMMARY OF RESULTS. Overall, my observations strongly confirm Geraghty's (1990) hypothesis that *R tends to be lost more and more "in proportion to distance from Western Oceanic"—that is, in the case of Vanuatu, in proportion to "southerliness" (Geraghty 1989:148). Geraghty's hypothesis was formulated on a broad geographical scale, encompassing languages from the Solomon Islands to Polynesia. His set of data was limited, in the case of Vanuatu, to just a few languages. Thanks to the information that has become available in the last two decades, it has been possible to propose finer-grained observations regarding the loss of *R.

Table 3 summarizes the results presented in the previous sections, and incorporates the items that will be discussed in the appendices.[33] The three major classes of examples are shaded.

I have examined 92 lexical items in total. Out of these, 35 (=38%) showed a neat divide between two areas: a northern area that retains *R, and a southern area that shows its loss. Crucially, the other major categories of table 3 are also compatible with this geographical pattern. For example, those words that have lost *R everywhere in Vanuatu have normally kept it in the Solomon Islands (Geraghty 1990).[34] The 17 relevant etyma therefore follow an isogloss that encompasses all NCV languages, and separates again a

32. Raga and Uripiv are cited from Ross, Pawley, and Osmond (2008:167), Araki from François (2002:317), and Nguna from Wheatley (1992:255).
33. The use of # in the numbering of examples refers to the appendices (see footnote 18).

northern area where *R is retained, from a southern area where it is lost. Symmetrically, when *R is retained everywhere in North-Central Vanuatu, it is always lost further south (3.3). Even the latter configuration can therefore be represented as a split between two areas: a northern one (= Solomons + NCV) conservative of *R, vs. a southern zone where *R is lost. Finally, some words (52)–(55) have retained *R in north Vanuatu, but are lost, or undocumented, for languages further south; even if these cases do not positively illustrate the same north-to-south tendency, at least they do not contradict it.

Except for one case (2.4), the number of cognate sets that support—or at least do not contradict—the north-to-south cline adds up to 91 out of 92 (=98.9%), which is a particularly strong correlation. Before I turn to the interpretation of this state of affairs, I propose to represent the facts on a map.

3.2 A MAP OF ISOGLOSSES. The loss vs. retention of *R can be represented visually by drawing each isogloss on the map of northern Vanuatu (map 3). Each circle represents a language that was cited in the paper. Map 3 represents the 15 distinct isoglosses discussed in the previous pages. Each of these lines includes not only the languages in this map, but also encompasses—as a general rule—other languages further south. The maximal geographical extension of these isoglosses is the topic of 3.3.

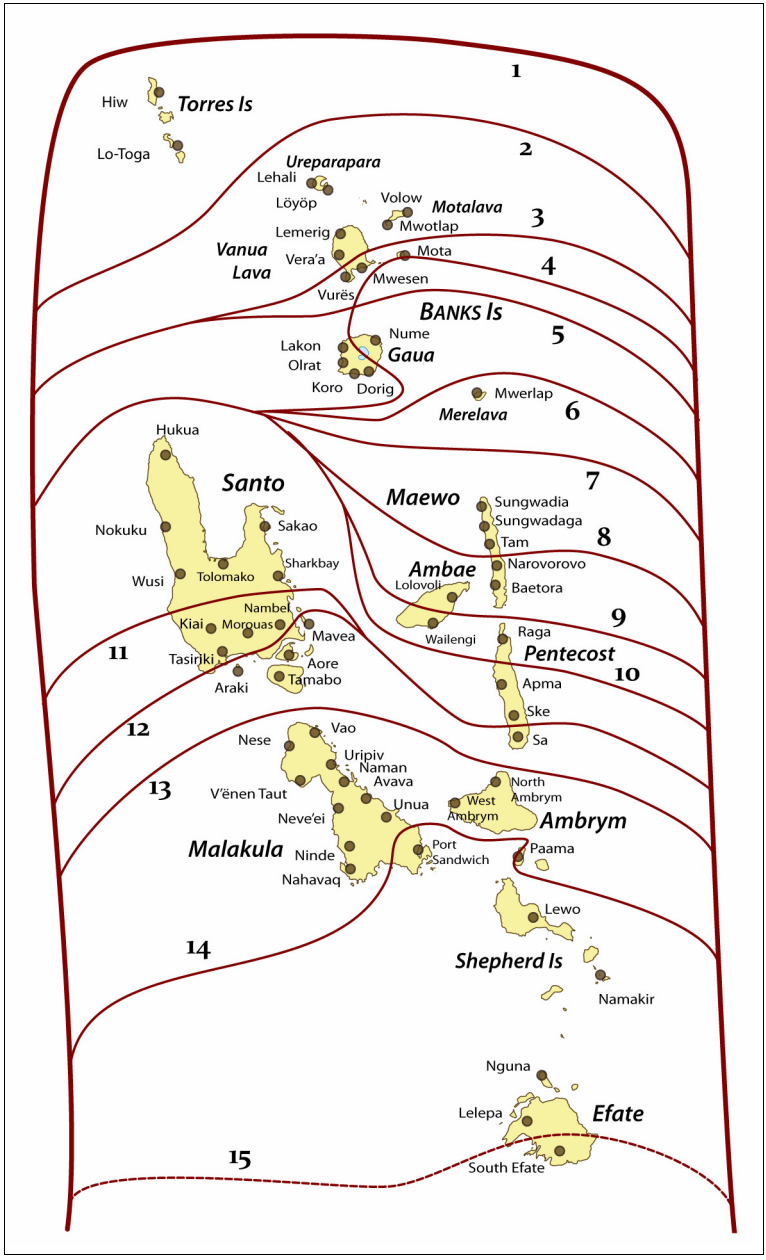
- {i1}: *R is present in Solomon Islands languages, but is lost throughout the North-Central Vanuatu linkage in a number of words (17 cognate sets, first row in table 3), e.g., ***baReko** 'breadfruit', ***piRu** 'fan palm', ***tapuRiq** 'conch shell'.
- {i2}: South of this line, all languages lose *R in ***kuRiap** 'dolphin'.
- {i3}: South of this line, all languages lose *R in ***[ka]ŋaRi** '*Canarium* sp.'
- {i4}: South of this line, all languages lose *R in ***quRis** '*Spondias*'.
- {i5}: South of this line, all languages lose *R in ***ñañoRap** 'yesterday', ***waRisa** '2 days from now', ***qana-waRisa** '2 days ago'. Possibly same line for ***puRe** 'Ipomoea sp.'
- {i6}: South of this line, all languages lose *R in ***kaRat** 'bite'.

TABLE 3. *R ETYMA SHOWING REFLEXES IN NORTH-CENTRAL VANUATU: A SUMMARY

Category	Examples	Number
Etyma losing *R in all NCV languages	(#8)–(#22), (#62), (#64)	17
Etyma losing *R in some NCV languages, showing patchy distribution		0
Etyma losing *R in some NCV languages, showing neat divide where *R is lost in north, retained in south	(56)	1
Etyma losing *R in some NCV languages, showing a neat divide where *R is retained in north, lost in south	(17)–(51)	35
Etyma only retained in some northern languages, where they retain *R; etymon is lost or undocumented further south	(52)–(55); (#60)–(#61); (#63)	7
Etyma retaining *R in all NCV languages	(#23)–(#54)	32
TOTAL: *R-etyma showing reflexes in NCV languages		**92**

34. This statement can be refined by observing the loss of *R in Vanikoro, the closest island to the north of the Torres Is: see appendix 3. Even though a handful of lexical items show *R-loss in both areas, I will argue that the evidence is too weak to argue against the possibility of chance.

MAP 3. PATTERNS OF RETENTION AND LOSS OF *R IN NORTH-CENTRAL VANUATU. (All isogloss boundaries divide the area along a north-to-south cline.)



- {i7}: South of this line, all languages lose *R in ***kuRita** 'squid', ***draRaq** 'blood'; ***takuRu** 'back', ***taRaq** 'cut'. Possibly same line for ***Ruap** 'rising tide', ***cakaRu** 'reef', ***d[r]aRaka** 'nutmeg', ***daweRu** 'coconut crab', ***kaRupe** 'ghost crab', ***tapoRa** '*Terminalia*', ***taRam-i** 'answer.
- {i8}: South of this line, all languages lose *R in ***suRi** 'bone'. Probably same line for ***paRu** 'hibiscus', ***paRage** '*Pangium*'.
- {i9}: South of this line, all languages lose *R in ***suRuq** 'fluid'.
- {i10}: South of this line, all languages lose *R in ***Ro?oti(k)** 'tie, bundle'. Possibly same line for ***(c,z)iRi(v)** 'grate, peel'.
- {i11}: South of this line, all languages lose *R in ***paRi** 'stingray'.
- {i12}: South of this line, all languages lose *R in ***(ma-)Rapun** 'haze'.
- {i13}: South of this line, all languages lose *R in ***kiRe** 'pandanus'. Possibly same line for ***b^{w}akaRe** 'pufferfish'.
- {i14}: South of this line, all languages lose *R in ***vaRasi** 'tread on'. Possibly same line for ***kaRi** 'bivalve'.
- {i15}: Southern limit of the North-Central Vanuatu linkage. South of this line, *R is lost in ***puRuk** 'cough'.

3.3 HOW FAR SOUTH? SOUTHERN VANUATU AND NEW CALEDONIA. The phenomenon illustrated by map 3 suggests that the languages of North Vanuatu are located on the northern fringe of a dialectological phenomenon that has its roots much further south. This raises the issue of how far south the isoglosses of *R-loss reach.

3.3.1 South Efate. The southernmost languages of the North-Central Vanuatu linkage are located on Efate (Clark 2009). Available data suggest that the north-to-south cline is respected all the way down to these: that is, languages of Efate have lost more *R than languages of Ambrym further north, which in turn have lost more than those of Santo, and so on. South Efate, which Clark (2009:56) considers to be the southernmost language of NCV, appears to retain and lose *R in exactly the same words as other NCV languages immediately to its north (for example, Lelepa and Nguna), judging by the data in Lynch (2000c:326) and Thieberger (2007).

In addition, whereas (#53) POc *puRuk 'cough' seems to have retained *R in all other NCV languages (Clark 2009:231), it is reflected in South Efate as /puk/ (Thieberger 2007):

(57) POc *puRuk > PNCV *vuru 'cough'
MTP wu**j**; MTA βu**r**u. … Nokuku wu**r**; Tamabo, Maŭea vu**r**u; Paamese hi**l**u; Lewo wu**r**i; ◆ South Efate puk

This suggests another isogloss {*i15*} similar to the ones drawn in map 3, separating again languages preserving *R from others further south losing it. On present evidence, South Efate is the most *R-less language within NCV.

3.3.2 Southern Vanuatu languages. According to Geraghty (1990:86), Southern Vanuatu languages "continue almost perfectly the Solomons-Northern Vanuatu trend of *R being lost from north to south." This observation is confirmed by additional data now available.

Table 4 compares the retention of *R in North-Central Vanuatu with relevant lexical reconstructions proposed by Lynch (2001) for Proto-Southern Vanuatu. For expository purposes, the whole Vanuatu archipelago is here divided into three arbitrary regions: Torres-Banks, the rest of North-Central Vanuatu, then Southern Vanuatu.

When any of these areas is consistent in retaining or losing *R, this appears respectively as "*R" or as "*Ø." When the isogloss of *R-loss cuts across one of these areas, this is represented as "*R | *Ø" within the same box; as we saw earlier, even these cases follow a consistent north-to-south cline. An empty box means I don't have the data; a box with "—" means that the etymon has been lost in the whole area.

This table provides ample evidence that the loss of *R in Southern Vanuatu languages affected (a) exactly the same lexical items as in islands further north, plus (b) other items specific to Southern Vanuatu. This can be formulated as two complementary tendencies:

(58) When *R is lost somewhere in Vanuatu, then it is also lost all the way to the south.

(59) When *R is retained somewhere in Vanuatu, then it is also retained all the way to the north.

These two strong tendencies admit only two exceptions—see the last rows of the table. Thus, whereas *paRi 'stingray' loses *R in languages much further north (isogloss {*i11*} in map 3), it retrieves it, as it were, in Sye *uvar* (Geraghty 1990:87, Lynch 2001:43). Likewise, Sye has *navruɣ* 'cough' (Crowley 2000:35), retaining *R where South Efate (*puk* < *puRuk) loses it. These cases of sporadic retention of *R in Southern Vanuatu languages (Lynch 2000a:59) only concern a small handful of lexical items, and can be compared with the cases of sporadic retention we saw in languages of Malakula. I will come back to this issue in the discussion below in 4.3.

TABLE 4. THE GEOGRAPHY OF *R-LOSS ACROSS VANUATU LANGUAGES

POC/PEOC etyma	**Torres-Banks**	**Other NCV**	**Southern Vanuatu**
*mawiRi	—	*R	*R
*RapiRapi – *yaRu – *bakuRa – *kaRaka – *tuRi – *kaRu – *vuRa	*R	*R	*R
*goRu – *kaRi – *baReko – *saRe	*R	*R	
*wakaR – *qaRa – *Ropok – *usuRi – *soRo(v) – *boRe	*R	*R	*Ø
*bwakaRe – *suRuq –*paRu –*suRi – *draRaq	*R	***R** | ***Ø**	*Ø
*kiRe	*R	***R** | ***Ø**	
*daweRu	*R	*Ø	
*kuRita – *kaRupe – *takuRu – *draRaq – *taRam(i) – *taRaq – *paRi- – *Rabia	*R	*Ø	*Ø
*paqoRu – *Ruqa – *uRat –	—	*Ø	*Ø
*ñaõoRap – *waRisa – *kaRat – *puRe – *quRis – *[ka]ŋaRi	***R** | ***Ø**	*Ø	*Ø
*tuRa – *boRok – *tapuRiq – *Rumwaq – *tuaRi – *piRaq – *meRaq	*Ø	*Ø	*Ø
*joRaga	*Ø	*Ø	
*Ruap	*R	*Ø	*R
*paRi – *puRuk – *suRi	*R	***R** | ***Ø**	*R

3.3.3 New Caledonia. The only published observations on *R-loss in New Caledonian languages are to be found in Geraghty (1989:149) and Ozanne-Rivierre (1992:205). Both note that *R was usually lost in "Proto-New Caledonian," and "sporadically retained."

I propose below a survey based on the published dictionaries of seven languages.[35] Cognacy judgments are notoriously difficult with languages of New Caledonia, especially from the south of Grande-Terre (see Grace 1990), so some of these comparisons should be viewed as tentative. Occasionally, I give protoforms in Proto-New Caledonian (PNC) as they are proposed in Geraghty (1989),[36] even when I have no access to modern reflexes. Numbers in pointed brackets refer to NCV data cited elsewhere in this paper.

(60) *R lost everywhere in Vanuatu and in New Caledonia:

⟨#11⟩ *Rumwaq 'house' >	IAI uma, NYE, NEL, ZUA mwa, CEM mwà, XRC mwâ
⟨#13⟩ *tuaRi 'long time ago' >	PNC *tuai
⟨#22⟩ *meRaq 'red' >	PNC *mia(ʔ) > NYE mia, XRC mîâ, ZUA mii

(61) *R lost somewhere in Vanuatu, and then all the way south to New Caledonia:

⟨23⟩ *kaRat 'bite' >	NYE yac, XRC kè, ZUA cacai
⟨18⟩ *[ka]ŋaRi '*Canarium*'>	NYE kaing
⟨27⟩ *^{n}raRaq 'blood' >	PNC *dra > IAI dra-, NEL da, CEM cèwè
⟨26⟩ *kuRita 'squid' >	PNC *kuita > NEL ciixa, ZUA ciia, XRC ketè
⟨24⟩ *paRi- 'reciprocal' >	NYE, NEL, ZUA pe-, CEM i-
⟨28⟩ *takuRu 'back' >	PNC *tau-
⟨51⟩ *taRutu '*Diodon*' >	IAI kaat
⟨41⟩ *kaRupe 'k.o. crab' >	NEL waup
⟨38⟩ *cakaRu 'reef' >	PNC *chau

(62) *R lost both in NCV and in New Caledonia, yet with sporadic retention in Southern Vanuatu:

⟨30⟩ *suRi 'bone' >	NYE, NEL, ZUA du, CEM dūu-
⟨33⟩ *paRi 'stingray' >	PNC *pai > IAI ve, NEL pa, NYE, ZUA pe, CEM pè, XRC pé
⟨37⟩ *Ruap 'rising tide' >	PNC *wap > NYE, NEL wap, XRC kwè

(63) *R retained everywhere in Vanuatu, lost in New Caledonia:

⟨#23⟩ *(y)aRu '*Casuarina*' >	IAI iwajoo, CEM óowíi, XRC wayu
⟨#26⟩ *boRe 'dream' >	IAI bwi, Drehu pue
⟨#26⟩ *tuRi 'pierce, sew' >	NMI thoi, XRC chii

35. From north to south: IAI= Iaai (Ozanne-Rivierre 1984), NEL= Nêlêmwa (Bril 2000), NYE= Nyelâyu (Ozanne-Rivierre 1998), NMI= Nemi (Haudricourt and Ozanne-Rivierre 1982), PAI= Paicî (Rivierre 1983), CEM= Cèmuhî (Rivierre 1994), XRC= Xârâcùù (Moyse-Faurie and Néchérö-Jorédié 1986). I am thankful to Claire Moyse for checking some of these forms with me, and to Isabelle Bril for sending data on Zuanga (ZUA).

36. For reasons I will not detail here, my PNC forms here differ from Geraghty's in two ways. I consider that *p was not lenited to *v at the level of PNC, and that diphthongs had not yet monophthongized. Hence 'stingray' is here reconstructed as PNC *pai, in contrast with Geraghty's form *ve.

⟨#32⟩ *kaRu 'swim' >	PNC *ghau > NYE yhao, NEL yha, ZUA zoo, XRC xê
⟨#47⟩ *[vo]m^{w}aRaki > '*Chalcophaps*'	NYE wêêk
⟨#27⟩ *soRo(v) 'call' >	PNC *cho
⟨#54⟩ *wakaR 'root' >	NYE wââ-, NEL waa-, ZUA we-, XRC kwèè-
*mawiRi 'left hand' >	PNC *mau > IAI me, NEL hma, NMI mo, PAI aè\mwü

(64) *R retained everywhere in Vanuatu and in New Caledonia:

⟨#43⟩ *saRe 'rip, tear' >	NEL da**l**e [but: etymon lost in SV]

(65) *R retained everywhere in NCV and in New Caledonia, yet showing sporadic loss in Southern Vanuatu:

⟨#48⟩ *suRi 'follow' >	NYE huu**r**, NEL huu**r**i, ZUA ho**r**e

(66) *R lost in (at least the southern parts of) Vanuatu, but retained at least sporadically in New Caledonia:

⟨10⟩ *piRaq '*Alocasia* taro' >	NEL, NMI, ZUA pia, *but* NYE pe**r**a, CEM pí**d**ú
⟨#12⟩ *tapuRiq 'triton' >	NEL kaawo, PAI tuū, *but* NMI daah**l**ook
⟨#17⟩ *joRaga '*fehi* banana' >	NYE, NMI daang, *but* PAI da**r**é
⟨36⟩ *paRa(s) 'tread on' >	NMI pe**r**a
⟨22⟩ *puRe '*Ipomoea pescaprae*' >	XRC nù**r**ù
⟨45⟩ *paRu 'hibiscus' >	NYE pa/xi, NMI pook, CEM pà/tì, XRC pe, *but* Drehu e**l**u, Nengone e**r**u
⟨29⟩ *taRaq 'cut' >	NMI cei, *but* IAI cha**l**aa, CEM ta**l**i

Table 5 combines these New Caledonian data with those of Vanuatu languages given in table 4 above. Torres-Banks languages are here lumped with other NCV languages. When *R was sporadically retained in a region, this is coded as '*Ø ~ *R' (except for Malakula, which was detailed above).

This table shows that the loss of *R is essentially concentrated in the southern areas of the Vanuatu–New Caledonia continuum. However, it is ambiguous as to where the center of maximal *R-loss is located. On the one hand, words like *mawiRi seem to lose their *R only in New Caledonia, suggesting the latter territory would be the most *R-less area in the region. On the other hand, the last rows of the table suggest that Southern Vanuatu may have lost *R when at least some New Caledonian languages sporadically retain it. There does not seem to be a simple answer to this question. The only conclusion is that the highest concentration of *R-loss is to be found somewhere south of Efate, either in Southern Vanuatu or New Caledonia, with a few lexical items sporadically preserving *R in both regions.

3.4 A NOTE ON CONDITIONING. Thanks to these observations based on new data, one may want to rework Geraghty's (1990:85) attempt at identifying possible tendencies in terms of phonetic conditioning.

Thus, Geraghty proposed that in initial position, *R "tends to be lost before *u*, and retained before *a*." However, the newly available data provide several counterexamples to this tendency: *R in *Ruap 'high tide' is retained in all Torres-Banks languages as well as in Southern Vanuatu; but it is lost in *Rapun 'haze' for most languages of Vanuatu, and in *Raka '*Pueraria* sp.' or *Rabia 'starch' for all of them.

His second hypothesis was that *R tends to be preserved between identical vowels. But the majority of Vanuatu languages—at least south of the Banks Islands—lose *R in words like *kaRat 'bite', *taRam(i) 'answer', *draRaq 'blood', *boRok 'pig', *suRuq 'juice', *takuRu 'back', and so on. This proposed tendency is therefore not strongly confirmed.

Other patterns, however, might be identified. It seems that *R tended to resist attrition slightly more when it was immediately followed by a high vowel that was word-final: see, for example, in table 5 the relatively high rates of retention of *R in *goRu, *biRibiRi, *yaRu, *tuRi, *kaRu, *paRi, *ʔusuRi, as well as (in map 3) *kaRi. But there are also counterexamples, such as *tuaRi, *paqoRu, *suRi, *[ka]ŋaRi, all of which show massive loss of *R.

Conversely, a possible environment favoring the loss of *R might be the presence of the postvelar *q, presumably a glottal stop /ʔ/, or a uvular /q/. Indeed, one finds massive attrition of *R in words like *meRaq, *Ruqa, *paqoRu, *quRis, *suRuq, *tapuRiq, *piRaq, *taRaq, and *draRaq. This tendency suffers only one exception: *qaRa 'fence', which preserves its *R all the way south to Efate (Clark 2009:74). Similarly, the presence

TABLE 5. *R-ETYMA ACROSS SOUTHERN OCEANIC LANGUAGES

POC/PEOC etyma	North Central Vanuatu	Southern Vanuatu	New Caledonian
*vaRo	***R** | ***Ø**		
*goRu	*R		
*RapiRapi – *biRibiRi – *bakuRa – *kaRaka	*R	*R	
*yaRu – *tuRi – *kaRu – *vuRa – *mawiRi	*R	*R	*Ø
*paRi – *Ruap	***R** | ***Ø**	*R	*Ø
*boRe – *[vo]m^{w}aRaki	*R		*Ø
*wakaR – *qaRa – *Ropok – *soRo(v) – *puRuk	*R	*Ø	*Ø
*kiRe – *kaRi	***R** | ***Ø**		
*suRi	***R** | ***Ø**	***Ø ~ *R**	*Ø
*kaRupe – *draRaq – *kuRita – *paRi- – *[ka]ŋaRi – *kaRat – *cakaRu	***R** | ***Ø**	*Ø	*Ø
*bwakaRe – *suRuq – *taRam(i) – *Rabia – *takuRu – *quRis – *nañoRap – *waRisa	***R** | ***Ø**	*Ø	
*baReko – *tuRa- – *paqoRu	*Ø	*Ø	
*Rumwaq – *meRaq – *Ruqa – *tuaRi – *uRat	*Ø	*Ø	*Ø
*tapuRiq – *piRaq	*Ø	*Ø	***Ø ~ *R**
*joRaga	*Ø		***Ø ~ *R**
*boRok	*Ø		
*paRu – *puRe – *taRaq	***R** | ***Ø**	*Ø	***Ø ~ *R**
*ʔusuRi	*R	*Ø	***Ø ~ *R**
*paRas	***R** | ***Ø**		*R
*saRe	*R		*R

of a velar may also have favored *R-loss, at least south of the Banks, as in *baReko, *boRok, *bʷakaRe, *kaRupe, *kaRat, *takuRu, *cakaRu, *kiRe, *kuRita, as well as *[ka]ŋaRi and others. But again, this is not absolute, as witnessed by the resistance of *R in *bakuRa, *kaRaka, *wakaR, *kaRi, *kaRu, and *goRu. What we see here is therefore a mere statistical tendency (stronger for *q, weaker for *k) suggesting that *R was prone to delete when it was in the vicinity of a velar or postvelar consonant. One might interpret this as a case of interference (see footnote 17) linked with the phonetic nature of *R. However, what exact process is at work here is unclear. Could it be that *R, after all, was velar [ɣ] or uvular [ʁ], and that these tendencies illustrate dissimilation? But an alternative interpretation might as well suggest that *R was coronal (see 1.1), and tended to be lost precisely when the adjacent syllable had a consonant with a back articulation. These questions are still speculative, and do not rest on sufficiently solid ground to be solved.

All things considered, the deletion of *R may have obeyed some tendencies, but none of these seemed to have been as regular as other instances of sound change in the region. Now, as I suggested in 1.3, the lack of a simple phonetic explanation might be seen, paradoxically, as a strength rather than a weakness. Precisely because the change *R > Ø affected the lexicon in an arbitrary way, the neat distribution of isoglosses across the archipelagoes cannot be attributed to mere chance or parallel innovation (see 4.4). This is precisely what makes the phenomenon of *R-loss a key for understanding the linguistic history of the region.

4. AN EARLY CASE OF LEXICAL DIFFUSION ACROSS THE SOUTHERN OCEANIC LINKAGE.

The geography of *R-loss described in the previous sections outlines a vast area consisting of the two archipelagoes of Vanuatu and New Caledonia—in other words, what Lynch (2000a) proposes to identify as the "Southern Oceanic (SO) linkage." For the vast majority of lexical items, a neat north-to-south cline can be observed, whereby languages lose *R in more and more words as one goes south, except for a few sporadic retentions here and there. One can think of two models that could account for the observed patterns of *R-loss in North Vanuatu: the migration model, and the diffusion model.

4.1 THE MIGRATION MODEL.

One way to interpret isoglosses, especially when they define neat groupings of languages, is in terms of successive migrations. Let us imagine, for the sake of discussion, the archipelago of Vanuatu and New Caledonia as an empty territory, being slowly settled by a population migrating south. The general north-to-south cline we have observed in the loss of *R could then be tentatively interpreted as a linguistic corollary of such a migration. Each step in this southward movement would have been followed by a pause, during which the dialect of the southernmost population, for some reason, would have lost *R in an additional number of lexical items.

Under this scenario, some of the lexical items represented by isogloss {*i1*} on map 3 could reflect innovations shared by "Proto-Southern Oceanic," and reflected in all modern SO languages: for example, *baReko > *baeko '*Artocarpus*'. Then isogloss {*i2*} (loss of *R in *kuRiap 'dolphin') could define a putative split between the two Torres languages vs. all the rest of SO, call it "Nuclear Southern Oceanic." The isoglosses

defined by line {*i7*} (loss of *R in *kuRita 'squid', and others) would define a further split between Torres-Banks and all the rest (see Tryon 1976:51). Further south, the loss of *R in words like *soRo(v) 'call' (table 5) would have taken place in "Proto-Southern Melanesian," the common ancestor of languages spoken south of Efate (see figure 1); and so on and so forth.

Should this migration scenario be confirmed by other sets of exclusively shared innovations, it could be represented as a branching family tree, in which each subgroup would correspond to a population departing from its northern relatives, and carrying along, in their migrations, an innovative speech variety characterized by the loss of *R in more lexical items than its immediate ancestor.

However, a number of problems would arise from such a representation. First, the classical family tree, if it is interpreted in terms of successive migrations, makes it difficult to accommodate pockets of conservatism, that is, the sporadic retention of *R in Malakula (in *taRaq 'cut', for example), in Southern Vanuatu (as in *paRi 'stingray'), or in New Caledonia (as in *piRaq '*Alocasia* taro'). For example, considering *piRaq loses its *R in virtually all of the 130+ SO languages, one might want to situate this loss at the time of their common ancestor, Proto-Southern Oceanic; however, the fact that it is exceptionally retained in Nyelâyu (Ozanne-Rivierre 1998) makes this assumption impossible—at least under the common understanding of the family-tree model. In each case like this one, one can only save the migration scenario by formulating ad hoc hypotheses, such as some borrowing from a non-SO language; or one may have to propose that the two forms *piRaq and *piaq coexisted in the speech of early settlers, and that the two variants percolated down the family tree until descendant languages eventually selected one of the two variants. And, of course, one would need to explain why 99 percent of modern languages ended up retaining the *R-less form, while one or two languages kept the *R. These are the sort of unlikely scenarios that the family-tree model, understood in terms of demographic splits, sometimes forces us to reconstruct. Conversely, such unexpected retention will become much easier to explain under the diffusion model that I shall advocate below (see 4.3).

The second reason why the migration model does not seem adequate here has to do with what we know of the settlement history of this part of Melanesia. As Pawley (1999:129) points out, the ideal phylogenetic tree corresponds best to a migratory scenario that involves pauses, allowing for each intermediate protolanguage to form over a period of time, and become stabilized enough before its speakers move on to settling other islands. Such a "pulse-and-pause" scenario is, for example, advocated by Pawley and Ross (1995) to account for the formation of Proto-Oceanic in the Bismarck archipelago, and indeed it is relevant to account for the formation of some clearly identifiable subgroups within Austronesian (Gray, Drummond, and Greenhill 2009). However, should such a scenario account for the distribution of *R in Vanuatu, this would imply that the fifteen isoglosses in map 3 correspond to as many successive stages in a very slow southward colonization.

Such a model would contradict archaeological findings, which show a rapid first settlement of Island Melanesia around 3,200 years ago, by the bearers of the "Lapita" civilization (Kirch 1997, Spriggs 1997). Just within the Vanuatu archipelago, sites

belonging to the Lapita cultural complex have been located in a number of distant islands: (a) Motalava Island in the Banks group (Shing et al. 2007); (b) Big Bay, north of Espiritu Santo (Bedford and Spriggs 2008:103); (c) Aore Island, south of Espiritu Santo (Galipaud and Vienne 2005); (d) islands off the northeast coast of Malakula (Bedford 2003); and (e) Teouma on the south coast of Efate (Bedford and Spriggs 2008:103). As far as Vanuatu is concerned, earliest radiocarbon dates for artifacts found in these Lapita sites revolve around 3150 BP (Bedford and Sand 2007:15), with earliest dates in Aore (south of Santo) and Efate. It is a well-established fact that, by 3000 BP, Lapita colonizers had settled most of Island Melanesia, all the way to New Caledonia and Fiji; this rapid colonization of Island Melanesia by the Lapita settlers involved no significant pauses (Green 2003, Bedford 2006, Pawley 2007b:24, Bedford and Spriggs 2008:97). Obviously, the slow settlement scenario, which would result if we interpreted the distribution of *R in Vanuatu in terms of sequential splits and migrations, is not easy to reconcile with this rapid expansion and settlement across Melanesia.

4.2 THE DIFFUSION MODEL. Both the linguistic and archaeological evidence make it necessary to propose a different scenario in order to reconstruct the whole linguistic history of the region, and in particular to account for the distribution of *R in Vanuatu.

Here is what Pawley (2007a:11) says about the linguistic aspects of the colonization of island Melanesia by bearers of the Lapita culture: "The rapid spread of Lapita from the Bismarcks to West Polynesia between 3200 and 2900 BP had a linguistic correlate. The speech of the Lapita colonists in the different island groups must have been relatively homogeneous, little differentiated from Proto Oceanic." The archipelagoes of Vanuatu and New Caledonia were thus colonized swiftly by a population that was relatively homogeneous. A linguistic correlate of this rapid dispersal is that the protolanguages we try to reconstruct as intermediate steps in the history of Oceanic languages (PNCV, Proto-Southern Oceanic, Proto-Central Pacific, and so on) all end up being very similar to what we reconstruct as Proto-Oceanic. This strongly suggests that no intermediate node had had the time to form during the process of settlement.

However, it is also true that a number of linguistic innovations are widespread in certain areas, for example, in North-Central Vanuatu (Clark 2009), or Southern Vanuatu (Lynch 2001). But due to the rapid dispersal of Lapita colonizers across the region, we know that these shared innovations cannot be understood as diagnostic evidence for "subgroups" in a genealogical sense, whereby modern languages would all descend from a single protolanguage spoken once by a compact community. Instead, when modern languages share a given innovation, this results historically from a process of in situ diffusion, via "horizontal" transmission (Enfield 2008), across populations who were already dispersed in space when this innovation began. Some of these postdispersal innovations expanded across entire dialect networks, while others were more limited in their spread.

Such a "wave" model (cf. Schmidt 1872) has been suggested for Vanuatu languages by Tryon (1996), and it is also implicit in the title of Clark (1985): "Groups, chains, clusters and waves." This is also the conception that underlies the concept of "linkage" (Ross 1988:8, Pawley and Ross 1995) in contrast with the family-tree notion of "subgroup." A linkage results from the in situ diversification of an earlier dialect network. Various studies

(for example, Geraghty 1983, Pawley 1999, Pawley and Ross 1995) discuss how linkage phenomena among Austronesian languages make it necessary to interpret the family-tree model with caution:

> When a group displays this kind of pattern of overlapping innovation sets, it is an "innovation-linked subgroup" [i.e., linkage–A.F.]. In this case, we infer that its members are descended from an earlier dialect network. Innovations occurred in various dialects, each spreading into neighbouring dialects, but not across the whole network. The crucial point about an innovation-linked subgroup is that its innovations give us no evidence of an exclusively shared protolanguage (Lynch, Ross, and Crowley 2002:92).

Thus, innovations shared between languages in a linkage may not allow us to conclude that they descend from a single ancestor. What they tell us is not about shared *ancestry* strictly speaking, but about shared *development* within a social network. Such observations can tell us a lot about ancestral social and linguistic networks in the early history of this archipelago.

The history of *R-loss is clearly an example of such a diffusion. Whereas the pulse-and-pause model would have suggested a consistent *southward* migration, instead the diffusional model I here propose identifies a *northward* direction for the spread of a particular innovation—the loss of *R—over already established speech communities. When that process of sound change began, the archipelagoes of Vanuatu and New Caledonia had already been colonized by a population speaking very similar dialects of Oceanic. Crucially, these initial dialects had evidently kept all (or most) instances of POc *R in their vocabularies. Later on, some time after this initial colonization, a new articulatory trend—the tendency to delete the liquid *R in some words—emerged somewhere south of Efate, whether in southern Vanuatu or New Caledonia. This new habit then began to diffuse to other islands, both northward (toward the Torres Is.) and southward (toward the southern areas of New Caledonia), through diffusion across a widespread dialect network.

4.3 THE STATUS OF EXCEPTIONS IN A DIFFUSION MODEL. I now briefly turn to the few exceptions we saw earlier. By exceptions, I here refer to the handful of languages—in Malakula, Southern Vanuatu, and New Caledonia—that show the unexpected preservation of *R in several words, in spite of other languages around them tending to lose the consonant in the same items.

At first glance, these few exceptions could be said to somewhat tarnish the otherwise very neat observations I made regarding the loss of *R. For example, the loss of *R in *paRi 'stingray' delineates a very coherent geographical area (*{i11}* on map 3) that runs from the middle of Santo all the way to the south of New Caledonia; all languages north of that line keep *R in *paRi, all the languages south lose it. Now, the fact that *R was retained in a few Southern Vanuatu languages (see 3.3.2) is puzzling, as it contradicts the north-to-south cline that is otherwise observed everywhere. Altogether, I have identified about 13 words (out of 92) that show this form of sporadic retention of *R within areas that otherwise lose it.

Paradoxically, these exceptions constitute a powerful argument in favor of my diffusional hypothesis. If the loss of *R had been interpreted in terms of shared innovations

inherited from specific nodes in a tree, then each case of exceptional retention would have provided a fatal counterexample to any tentative reconstruction (see 4.1). This is not the case in a diffusional model. What we have here is, for each word, a vast area that was massively affected by the horizontal spread of an innovation across a dialect network; and simply, here and there, residual pockets of conservatism, or *relic areas*. The imperfect distribution of innovations across dialect chains or networks is a typical feature of linkages—reflecting the post-dispersal horizontal spread of innovations across a dialect network (Ross 1988:8)—in contrast with subgroups, which reflect the vertical inheritance from a predispersal common ancestor. These relic areas are thus the cornerstone that makes it necessary to interpret the loss of *R as a case of postdispersal diffusion.

Interestingly, the loss of *R in various lexical items is not the only piece of evidence that supports the identification of Southern Oceanic (or of NCV for that matter) as a *linkage* rather than a subgroup. Another conspicuous example is the treatment of POc *q, the "postvelar" stop. On the one hand, the consonant was lost systematically, and in all positions, in almost all of the 94 NCV languages—the sort of shared innovation it would be tempting to assign to Proto-NCV, the common ancestor of all these languages. However, this postvelar stop is still reflected as a glottal stop /ʔ/ in just one language, Namakir (Clark 2009:17), and has left traces in some Malakula languages (Lynch 2009a); as for New Caledonia, *q is sometimes lost, and sometimes preserved as a velar (Ozanne-Rivierre 1992:194). This situation, which combines a massively shared innovation (*q > *ʔ > Ø) with a few pockets of conservatism, is again best understood as a case of postdispersal horizontal diffusion across a linkage, rather than an innovation that took place in any identifiable node of a phylogenetic tree.

The same demonstration could be carried out for a great number of widely but imperfectly distributed features within the Southern Oceanic linkage, for example, loss of word-final consonants (Ozanne-Rivierre 1995, Lynch 2005); lenition of voiceless stops (like *k > ɣ, *p > *v*) and other cases of sound change;[37] possessive classifiers; numeral systems (Lynch 2009b); and a number of others. But this would take us beyond the scope of this paper. Suffice it to say, the distribution of *R reflexes is just one among many linguistic features that make it difficult to apply the family-tree model in this part of Melanesia. The distribution of many (if not all) of the innovations that are shared among Vanuatu languages is better explained by a wave-model approach than by a model based on hierarchized and nested subgroups (see also Pawley 1999; Garrett 2006; Heggarty, Maguire, and McMahon 2010).

4.4 PARALLEL INNOVATION? A possible objection to the unity of the phenomenon here described would be to underline the fact that *R already shows a tendency to disappear in other parts of the Oceanic family (see 1.2). As Geraghty (1990:83) points out, Proto–Malayo-Polynesian (PMP) *R inexplicably disappeared in three words as early as Proto-Oceanic (as with PMP *maRi > POc *mai 'come'), and in a great number of words in Micronesian and Central Pacific languages. To these areas, I will add the little-described Temotu subgroup, in particular the languages of Vanikoro, which also have lost *R in at least ten words (appendix 3).

37. Lynch and Brotchie (2010) provide a similar discussion showing how the apicolabial shift [e.g., *m > *m̈ > *n*] must have spread historically across a dialect network in northern Malakula.

This tendency for *R to be lost cannot be explained easily, due to the unclear phonetic status of the consonant, and the general absence of conditioning factors. But if this phoneme was unstable anyway, then this argument might seem to weaken the idea that *R-loss diffused all the way from New Caledonia to northern Vanuatu as a unitary phenomenon. An alternative hypothesis would resort to parallel innovation rather than diffusion (see Geraghty 1990:90).

The way to tackle this problem is to observe exactly which words were affected. From reading Geraghty's (1990:88) data, it appears that other language groups—say, Micronesian languages—lost *R sometimes in the same words as Vanuatu languages, and sometimes in different words. Although the general tendency of dropping *R is found in the two distant archipelagoes, few individual lexical isoglosses are actually shared across them. In this case, parallel innovation is probably the likeliest interpretation. This will also be my conclusion in the case of Vanikoro, the island of the Solomon Islands that is closest to northern Vanuatu (see appendix 3): while *R was lost there sometimes in the same lexical items as in Vanuatu, their number is too small to support a solid hypothesis in terms of diffusion. The same could be said about Central Pacific languages.

However, the situation is different in the case of Southern Oceanic languages. The waves we need to account for are not just the spread of a general articulatory habit (the habit of dropping *R in all environments), but the spread of 91 specific individual lexical innovations showing unpredictable loss of *R. We need to explain how the habit of saying *pai instead of *paRi for 'stingray' spread northward across the social network of Vanuatu, before eventually stopping its course at isogloss *{i11}*; how the habit of saying *ʔaŋai instead of *ʔaŋaRi for 'almond' spread all the way to isogloss *{i3}*; and so on. The existence of a continuum from Vanuatu across to New Caledonia for a great number of lexical items (see 3.3.3) is not likely to be accidental. Each of these lexical items constitutes a shared innovation, and their accumulation provides strong support to the diffusional hypothesis.

4.5 DIFFUSION ACROSS DIALECTS, DIFFUSION THROUGH THE LEXICON. Finally, it is reasonable to assume that the spread of *R-loss did not take place between separate languages, but across members of a vast dialect network at an early date when they were still close to each other. The two arguments for this conclusion are, first, the degree of mutual intelligibility that is necessary for such an irregular innovation to diffuse across dialects (this section); and second, the relative chronology of *R-loss with respect to its merger with *r (see 4.6).

The high degree of language diversification known in modern Vanuatu and New Caledonia implies that these languages ceased to be mutually intelligible many centuries ago. Over the generations, what started 3,000 years ago as a far-flung network of small voyaging communities maintaining some form of social and linguistic unity slowly crumbled into smaller social networks. As local populations grew and developed economic self-sufficiency, they relied less and less upon their distant relationships. Eventually, the social bonds between northern islanders and their southern cousins were reduced to a minimum. This tendency toward the breakup of earlier networks is observed by archaeologists: "there is regional diversification in ceramic traditions right across Vanuatu

soon after Lapita … [and] these traditions can be divided into a number of regions: southern, central, northern, and far northern Vanuatu (Banks and Torres Islands)" (Bedford and Spriggs 2008:107).

Pawley underlines the linguistic effects of this historical tendency toward internal diversification: "After the first phase of colonisation, the archaeological and linguistic record indicates that in the Southern Melanesian archipelagos a sequence of demographic and cultural changes occurred which led to weakening or loss of communication between distant sister communities.… Most linguistic innovations spread only short distances and the speech traditions of distant communities diverged" (Pawley 2007a:21). This breakup of earlier social networks resulted in increased cultural and linguistic fragmentation, first in the form of dialect differentiation within a once more homogeneous dialect chain, and later in the form of distinct languages constantly diverging from each other (François forthcoming).

Now, could it be the case that the loss of *R took place late in this process of linguistic fragmentation, at a time when neighboring languages had already lost mutual intelligibility? This is highly unlikely, for a number of reasons. First, this form of sound change seldom cuts across language boundaries. Some examples of contact-induced sound change spreading across languages have been reported elsewhere, such as the change from [r] to [ʁ] in middle Europe, historically attested in Germanic and Romance languages in contact (Chambers and Trudgill 1998:170–75). However, this sound change concerned the spread of a socially marked articulatory habit affecting the realization of one phoneme across the whole lexicon of each speech variety, regardless of which individual lexemes it occurred in. The case of *R-loss is very different, because it affected only some lexical items and not others, in ways consistent across adjacent regions (see 4.4), with no obvious conditioning principles (see 3.4).

Because the sound change *R > Ø only affected certain words and not others, it constitutes a case of "lexical diffusion" (Wang 1977, Labov 1994:421ff., Bybee 2002), that is, a change that developed word by word over a long period of time, and affected the lexicon gradually. Note that the use of the term *diffusion* here does not refer to the diffusion in space—from dialect to dialect—but the diffusion of the sound change within the lexicon of each dialect. Bybee (2002) showed that gradual diffusion is typically correlated with word frequency in discourse, a hypothesis that is difficult to test in our case. Labov (1994:542), and after him Krishnamurti (1998:195), contrast *lexical diffusion* ("the word as a unit of change") with the *regular sound change* of the Neogrammarians ("the phoneme as a unit of change"), and argue that both occur in language change. The process of gradual lexical diffusion sometimes stops halfway through the lexicon, but sometimes goes through to completion in some languages, in which case the sound change appears, at least retrospectively, to be "regular" (Krishnamurti 1998). The loss of *R across the Southern Oceanic dialect network corresponds to such a case of lexical diffusion that proceeded word by word, spreading through the lexicon at the same time as it was spreading across dialects. In some dialects of New Caledonia, the change *R > Ø eventually swept through the whole vocabulary, whereas further north its diffusion (both dialectal and lexical) was stopped halfway; I will propose an explanation for this in the next section.

The spread of such irregular sound change, because it affects only certain items and not others, is unlikely to have taken place across languages after they had already lost mutual intelligibility. In contrast, it makes much more sense if we set ourselves at an early time in the settlement of these territories, when the archipelagoes were integrated into vast networks of social exchange and relative linguistic unity.

Strictly speaking, the spread of these lexical innovations should not be understood as cases of lexical *borrowing*, at least not in the traditional sense of borrowing that involves separate languages. The replacement of *paRi by *pai presumably consisted in the alternation of two variants of the same word within (adjacent varieties of) the same language, rather than as the adoption of a foreign form from another language; the social and cognitive processes involved in the two situations are different (Hock 1991:388, Hock and Joseph 1996:342). Indeed, sound change will typically spread from one speaker to another if they view themselves as speaking the "same language" (François 2001:16, Enfield 2003:20), or at least mutually intelligible dialects. This is a situation when speakers will be readily inclined to imitate each other's speech habits, as they identify themselves with a certain social group (Le Page and Tabouret-Keller 1985; Labov 1963, 1994, 2001). By adopting the new speech habit from their peers, these speakers de facto perpetuate the social coherence of their speech community.

4.6 TWO SOUND CHANGES IN COMPETITION. Finally, another reason that the loss of *R is necessarily a very ancient phenomenon is the relative chronology that needs to be posited with regard to its merger with other consonants.

One of the properties of *R in Oceanic languages is that its modern reflexes all result from a merger with another liquid, *r for some (proto)languages, *l for others (Ross and Næss 2007:472). This other liquid always shows more stability than *R: for example, we have seen in 2.2.1 that *r is reflected regularly in all Vanuatu languages, and does not show the same instability as *R.

Clearly, the spread of *R-loss must have taken place early in the history of Oceanic languages, at a time when this consonant was still distinct from other liquids. Considering how several subgroups of Oceanic are identified—along with other innovations—by what consonant *R was merged with, this also means that the loss of *R took place at a very early stage in the unity of Southern Oceanic, before it began breaking down into regional subgroups. First the loss of *R in some words, and soon after it, its merger with *r, must both have taken place, in this order, at a time of relative linguistic unity. This is not necessarily to say that the whole Southern Oceanic archipelago, at the time when *R-loss diffused, spoke a single language with no variation; on the contrary, the difference in impact of *R-loss is precisely an example of such dialectal variation. Rather, the idea is that the region constituted a vast dialect chain à la Fiji (see Geraghty 1983), in which at least adjacent communalects had maintained enough mutual intelligibility with each other to allow for processes of sound change and lexical diffusion, which typically take place among speakers of the same language, to spread across the network.

All these observations suggest a plausible historical scenario. The loss of *R was a sound change that emerged somewhere in southern Vanuatu or New Caledonia, and began spreading across these two archipelagoes. As a change in progress, it was gradual in two

ways: on the one hand, it impacted various dialects in varying degrees; on the other hand, it only affected a subset of their lexicon, at least during the initial stages of the change. It was in the process of extending its effects across the whole lexicon when a *second* change took place, namely the merger of *R with *r. As this new merging tendency affecting *R quickly diffused across entire dialect networks, it competed with, and eventually bypassed, the tendency to its deletion, before the latter could affect all *R words. As a result, each dialect's vocabulary was suddenly frozen in its state at the time of the merger; some words had already lost their *R, but others had not, and were now becoming immune to the change.

This is how we are today able to observe the various isoglosses of *R-loss for individual words, as their spread was suddenly stopped halfway across the archipelago, and halfway across each dialect's vocabulary. Just as some glaciers keep memories of the climate in prehistoric times, so the geographical patterns of *R-loss provide us with a unique snapshot of the extension and shape of early social networks in this part of Melanesia, in the very first centuries of its settlement by Oceanic speakers.

5. CONCLUSION: THE FOSSILIZATION OF ANCIENT SOCIAL NETWORKS. This study of *R-loss still leaves a number of problems open for future research. What was the precise phonetic nature of *R in Proto-Oceanic? What was the motivation for its loss in some lexical items and not others? Was it phonetic conditioning? Was it frequency in discourse? Why did the same sound change take place in other Oceanic subgroups (Micronesian, Central Pacific, Temotu), albeit in distinct words?

This being said, the analyses presented in this paper already provide us with a number of insights that reach beyond just the fate of a single consonant. They identify a tangible marker that can help us understand the early linguistic history of this region. The distribution of *R in Vanuatu and New Caledonia portrays a vast archipelago at a crucial time of its history—between its early colonization by speakers of Proto-Oceanic, and the breakup of its social and linguistic network into a mosaic of mutually unintelligible languages.

A first important result of these analyses was to demonstrate that the whole archipelago of Vanuatu, despite its received division (cf. Ross, Pawley, and Osmond 1998:6) between the two linkages "North-Central Vanuatu" and "Southern Vanuatu," really formed a continuous dialect network for some time after its initial colonization. A similar link can be safely proposed between Vanuatu as a whole and New Caledonia. The case of *R thus adds to the evidence proposed by Lynch (2000) to identify "Southern Oceanic" as a linkage.

A second important finding from charting the *R-loss isoglosses is their relatively tidy distribution across the Vanuatu archipelago, in the form of a gradual cline ranging from north-to-south (map 3). I have proposed in 4.6 that these isoglosses represent a fossilized picture of the geography of dialects at the time of the merger between *R and *r, which put an end to the attrition of *R. Crucially, the orderly gradation of *R-loss isoglosses, which has been preserved until modern times, suggests that the spatial distribution of modern languages in Vanuatu essentially continues the geography of these early dialects, with no major interisland migrations posterior to that merger. Despite occasional oral histories of more recent movements on the local scale, the tie between each popula-

tion—or at least each language—and its territory thus appears to have been extremely stable in time, ever since the Vanuatu archipelago was initially populated.[38]

Finally, a third conclusion follows from the latter two observations. If the speech communities of Vanuatu remained a vast dialect chain for some time, and if modern languages continue directly the dialects spoken then in the same area, then this implies that the isoglosses observable today tell us something about the specific social networks that then linked islands together. For example, the band of territory located between lines *{i9}* and *{i11}* on map 3 suggests that the northern half of Santo at some point formed a highly coherent network with southwest Ambae, as well as with the northern part of Pentecost; but it had limited contact with Gaua, despite similar geographical distances. Rather than any direct Gaua–Santo connection, the social and linguistic ties between the Banks group and islands further south evidently followed a route via Merelava and Maewo. While such conclusions emerge from the reflexes of *R, they are also confirmed by what is otherwise known of these languages. Besides, they could be usefully complemented by further dialectological studies on various phonological, morphological, or lexical innovations.

Such future studies might either confirm the geographical patterns outlined here, thereby strengthening the case for these regional networks in ancient times, or might point to different groupings, due to changes in community relations during the course of history. Ideally, a representational system should be designed to capture these diffusional areas and their historical dynamics in a more accurate way than the misleading family tree, whose predictive power is ultimately grounded in innovation-defined subgroups and demographic splits. Finally, dialectological studies would profit from being intertwined with ethnographic and archaeological research, in the attempt to reconstruct the history of early social networks in this part of the world.

APPENDIX 1. PRIMARY DATA FROM THE TORRES-BANKS LANGUAGES

While the primary data underlying this study were usually cited in the body of the paper, sometimes the level of detail was not warranted by the demonstration, and could have hampered the reader's comfort. When the data were already published elsewhere, appropriate references were given. However, the primary data collected by the author in northern Vanuatu are published nowhere else, and are therefore grouped under the present appendix.

When Torres and Banks languages reflect regularly an etymon that has been reconstructed as PNCV by Clark (2009), I indicate the PNCV form. Conversely, when the form or its meaning are specific to the Torres-Banks area, I sometimes propose a reconstruction at the level of Proto–Torres-Banks (PTB). Rather than a protolanguage strictly speaking, the latter is better understood as a convenient way to encapsulate a number of innovations that have spread across the Torres-Banks languages through their history, mostly via diffusion (see 4.3).

1. REFLEXES OF *r

These few examples illustrate the regular reflexes of *r, the other rhotic of the protolanguage. See the discussion in 2.2.1.

(#1) POC *barapu > PTB *(ba)baravu 'long; tall':

38. Admittedly, while this point can be safely argued for Vanuatu, it cannot be demonstrated for New Caledonia, at least not based on the distribution of *R.

HIW pəᵍʟɔ; LTG pərɔ; LYP papjø; LHI pjɒ; LMG pœrœβ; VRA ᵐbaraβ; MTA paparau; DRG ᵐbuᵐbra; OLR pɒpra; MRL ᵐbʊrʊ ~ ᵐbɒmrʊ

(#2) POC *kara(s) > PNCV *karasi 'scrape (coconut)':
HIW ᵍʟi̯e; LTG ɣirɛ; LHI ɣɪjæh; LYP ɣɪjɪs; VLW ɣɛjɛh; MTP ɣɛjɛh; LMG ɣɛrɛs; VRA ɣaras; MTA ɣoras

(#3) PNCV *maraya 'moray eel':
HIW meᵍʟi̯e; LTG mərɨ; LYP n-mæjæ; VLW n-maja; MTP na-mja; LMG n-mara; VRA mɛriɛ; VRS marɪ; MSN mɛrɛ; MTA marea; DRG mrɛ; KRO mɛrɛ; OLR mɛrɛ; LKN marɛ; MRL nɛ-mɛrɪ

(#4) POC *raqup > PNCV *raʔu 'Dragon plum, *Dracontomelon vitiense*':
HIW ᵍʟɔ; LTG rɔ; MTP nɛ-jɛ; LMG wɔ\rɔ; VRA wɛ\rɛ; VRS wɔ\rɔ; MSN wɔ\rɔ; MTA rau; NUM wɛ\ra; DRG wa\ra; LKN ra

(#5) POC *rarap > PNCV *rara(vi) '*Erythrina indica*':
HIW ᵍʟaɣə (dissimil.); LTG rarə; MTP na-jaj; VRA raraβ; VRS rɛrɛβ; MSN rar; MTA rara ~ raraβ; NUM rɛrɛβ; DRG rra:β; LKN ræræβ

(#6) POC *rua > PNCV *rua 'two':
HIW βi-ᵍʟə; LTG bə-ruə; LHI bɪ-jɔ; LYP βʊ-jʊ; VLW βʊ-jʊ; MTP βʊ-jʊ; LMG βʊ-ru; VRA βʊ-ruʊ; VRS ni-rʊ; MSN ni-rʊ; MTA ni-rua; NUM ni-ru; DRG s-rʊ; KRO i-rʊ; OLR ni-rʊ; LKN ni-rʊ; MRL i-rʊ

(#7) PNCV *zara 'village clearing, central area in village':
HIW saᵍʟə; LTG sarə; LHI n-saj; LYP n-saj; LMG n-sɒr; VRA sara; VRS sar; MSN sar; MTA sara; NUM sar; DRG sar; KRO sar; OLR saj; LKN sa:; MRL na-sar

2. *R LOST EVERYWHERE IN NCV

I here provide only data from the Torres and Banks languages. For other NCV languages, see Clark (2009). See also the discussion in 2.3.1.

(#8) POC/PEOC *baReko > PNCV *baeko 'breadfruit, *Artocarpus altilis*':
HIW pɪɣ; LTG pɛɣ; LHI pæ; LYP n-pi͡ɛ; VLW n-ᵐbɛɣ; MTP nɛ-ᵐbɛɣ; LMG n-pɛɣ; VRA ᵐbiɛɣ; VRS ᵐbiɪɣ; MSN pɛx; OLR pɛː; LKN pɛɣ

(#9) POC *boRok > PEOC *b$^{(w)}$oRo > PNCV *boe 'pig':
HIW, LTG kʷo 'castrated pig'; LYP n-k͡pʷɔ; VLW nɔ-ᵑg͡bʷɔ; MTP nɔ-k͡pʷɔ; LMG k͡pʷœ; VRA k͡pʷʊ; VRS k͡pʷʊ; MSN k͡pʷɔ; MTA k͡pʷoe; NUM k͡pʷɔ; DRG k͡pʷɔ; KRO k͡pʷɔ; OLR k͡pʷɔ; LKN k͡pʷɔ

(#10) POC *piRaq > PEOC *viRa > PNCV *via '*Alocasia* taro':
HIW βi̯ə; LTG βiə; LHI n-βɛ; LYP n-βi͡ɛ; VLW nɪ-βɪ; MTP nɪ-βɪ; LMG n-βɪ; VRA βiɪ; VRS βɪ; MSN βɪ; MTA βia; NUM wɔ\βi; DRG βɪ; OLR βɪ; LKN βɪ; MRL nɛ-βɛ͡a. [Raga via; Araki via; Nguna na-via]

(#11) POC *Rumʷaq > PNCV *yumʷa 'house':[39]
HIW ɪŋʷə; LTG ə̯ŋʷə; LHI n-ɛŋ; LYP n-i͡ɛŋ; VLW n-ɪŋ͡mʷ; MTP n-ɪŋ͡mʷ; VRA n\iŋ͡mʷɪ; MSN ɪŋ͡mʷ; MTA iŋ͡mʷa; LKN uŋ͡mʷæ; MRL n-ɛ͡aŋ

(#12) POC *tapuRiq > PEOC *tavuRi > PNCV *tavui > PTB *tauwe 'conch shell, *Charonia tritonis*':
HIW təjə; LTG təwə; LHI n-tɔw; LYP n-tʊw; MTP na-tʊ; LMG n-ʔʊw; VRS tʊw; MSN tʊw; MTA tawe; DRG tʊw; OLR tʊw; LKN taʊ; MRL nɜ-tʊ

(#13) POC/PEOC *tuaRi > PNCV *tuai > PTB *tuai '[ADJ] old, ancient; [ADV] long ago':
HIW ᵍʟɔ\sse; LTG sɛ; LHI ttæ; LYP ʧø; MTP n\ɪtɛ; VRS ɣü\tüœ; MTA tuai

(#14) POC *tuRaŋ 'companion'> PEOC *tuRa- 'same-sex sibling'> PNCV *tua- 'sibling, friend':[40] > PTB *tua- 'companion, partner; other':

39. Rather than being a direct reflex of *R, the glide *y in *yumʷa probably developed as a way to strengthen the word-initial syllable onset. Note also that Lewo has *yumʷa*, Nguna has *sumʷa* (Clark 2009:236), and Proto-Southern Vanuatu has *n-iumʷaq.

40. Logically, the meaning 'companion' should also be reconstructed for the intermediate stages PEOC and PNCV.

MTP i\tan ‘the other one’; LMG ʔɒwɒ-n ‘the other one’; VRA ʔɔ-; VRS tœ-; MSN tɔ-; MTA tua-; NUM ta-; DRG tʊ-; LKN tɪ-; MRL tɔ-

(#15) POC *tuRaŋ ‘companion’> PEOC *tuRa- ‘same-sex sibling’> PNCV *tua- ‘sibling, friend’ > PTB *tuatua- ‘opposite-sex sibling’:
HIW tɯtβɔ-; LTG səse-; LHI tɛtɒ-; MTP tɪtɛ-; LMG ʔœwœ-; VRA ʔɔwɔ-; VRS tütüœ-; MSN tutuɔ-; MTA tutua-; DRG tuta:-; KRO tɛtɛ̯a-; OLR tati-; LKN tata-k; MRL tɛtɔ-

(#16) POC/PEOC *Raka > PNCV *aka ‘k.o. vine, *Pueraria* sp.’:
MTP n-aɣ; … Raga aga (Ross 2008:274)

(#17) POC *joRaga > PNCV *zoaga ‘*fehi* banana’:
DRG sʊak; KRO sʊak; LKN kpʷɪtɪ\sɔak; cf. Raga hoaga (Ross 2008:278)

Some terms are also members of etymological doublets (see discussion in 2.2.3):

(#18) POC *tabiRa ‘bowl’ > PNCV *tabia ‘wooden dish where pudding is pounded’:
HIW təpi̯ə; LTG təpiə; VLW n-tᵐbɪ; MTP na-tᵐbɪ; LMG n-ʔipi; VRA ʔiᵐbiɪ; VRS taᵐbɪ; MSN tɪpɪ; MTA tapia; NUM tiᵐbi; DRG tᵐbɪ; OLR tɪpɪ; LKN tapɪ; MRL taᵐbɛ̯a. Cf. doublet *taᵐbera (#60).

(#19) POC *kaRat ‘bite’ > PNCV *kati ‘chew (esp. kava)’:
HIW ɣɔt; LTG ɣɛt; LYP ɣɪt; MTA ɣat. Cf. doublet *kaRat (23).

(#20) POC *Rabia > PNCV *abia ‘starch, esp. sago starch’:
HIW epi̯ə ~ pi̯ə; LTG epiə; MTA pia-. Cf. doublet *rava (55).

(#21) POC *meRaq ‘red’> PNCV *mea ‘red pigment’:
HIW mi̯e; LTG mi; MTP nɛ-mɛ; VRS mɪ; MTA mea ‘red earth’; DRG mɛ; LKN mɛ

(#22) POC *meRaq ‘red’> PNCV *memea ‘red’:
HIW mi̯e; LTG məmi; LHI ɛmmæ; VRA mimiɛ; VRS mamɪ; MSN mɛmɛ; MTA memea; NUM mmɛ; DRG mmɛ; OLR mɛmɛ; LKN mɛmɛ; MRL mmɪ; … Lolovoli memea. Cf. doublets *mea (#21) and *mera (53).

3. *R RETAINED EVERYWHERE IN NCV.

Some words have retained *R everywhere among NCV languages, and later merged them with *r. This situation is summarized by Clark (2009) by using *r, rather than *R, in his PNCV etyma. See the discussion in 2.3.2.

(#23) POC *(y)aRu > PNCV *yaru ‘*Casuarina equisetifolia*’:
HIW ɔᵍʟ; LTG ɔr; LHI n-nɒj; LYP n-ɪj; VLW n-ɛj; MTP n-ɛj; LMG n-ɪr; VRA nɛr; VRS œr; MSN ɔr; MTA aru. […Araki vi-aru; Uripiv n-ur; Namakir ne-ar…]

(#24) PEOc *bakuRa > PNCV *bakura ‘tamanu, *Calophyllum* sp.’:
HIW pʉwɯᵍʟə; LTG pʉɣʉrə; LYP n-pʊɣʊj; MTP na-mwʊj; VRA ᵐbuɣurʊ; VRS ᵐbʊɣʊr; MSN pʊɣʊr; MTA pawura; NUM ᵐbuwur; DRG ᵐbɣʊr; LKN pawʊ:

(#25) POC *biRibiRi > PNCV *biribiri ‘k.o. shore tree, *Hernandia nymphaefolia*’:
HIW piᵍʟpiᵍʟ; LTG pərpir; MTP ni-ᵐbijᵐbij; VRA ᵐbirᵐbir; VRS ᵐbirᵐbir; MSN pirpir; MTA pirpir; NUM ᵐbiriᵐbir; DRG ᵐbirᵐbir

(#26) POC *boRe > PNCV *bore ‘dream’:
HIW kʷoᵍʟ, kʷᵍʟe; LTG kʷor; LHI kʷɛjkʷɛj; LYP k͡pʷɔjk͡pʷɔj; MTP k͡pʷɔjk͡pʷɔj; LMG k͡pʷœr; VRA k͡pʷʊr; VRS k͡pʷʊrk͡pʷʊr; MTA k͡pʷore; NUM k͡pʷɔrk͡pʷɔr; DRG k͡pʷɔr; LKN k͡pʷɔ:; MRL kʷœr

(#27) PEOc *(cz)oRo-v ‘call’ > PNCV *soro(vi) ‘snort, grunt’:
MTP nɔ-sɔj ‘cricket’; VRA sʊr ‘cricket’; VRS wʊ\sʊr ‘cricket’; MTA soro ‘make a droning noise like a tree-cricket’

(#28) POC *goRu(s) > PNCV *goru ‘(vegetation) dry’ > PTB *goru ‘(vegetation) dry; a biscuit made of dried breadfruit’:
HIW kɵᵍʟ; VLW nɪ-ᵑgɪj; MTP nɪ-kɪj; LMG kør; VRS kør; MSN kʊr; MTA kor; DRG kʊr; MRL kər

(#29) POC *tuRi ‘sew’ > PNCV *[t,d]uru ‘pierce, sew’ > PTB *duru ‘pierce, drill; make a hole’:
LTG ʈʉ**r**; MTP ⁿdi**j**; VRS ⁿdü**r**; MSN nu**r**; MTA nu**r**

(#30) PEOC *kaRaka > PNCV *kara[ka] ‘climb’:
HIW βɔ\ᵍʟɛᵍʟa**j**ə; LTG ɣəra**j**ə; MTP ɣa**j** ‘spread’; MTA ɣa**r**a ‘spread’

(#31) POC *kaRi(q)a > PNCV *[k,g]aria ‘*Cordyline terminalis*’:
HIW ti-ᵍʟeɣi̯ə (met.); LTG hɔ-ɣria; LYP n-ⁿdø-ɣa**j**ɛ; LMG tœ-ɣi**r**i; VRA ɣi**r**iɪ; VRS ⁿda-ɣa**r**ɪ; MTA ka**r**ia; NUM ⁿdɔ-ki**r**i; DRG k**r**ɪ; LKN (kæhræ)

(#32) POC *kaRu > PNCV *karu ‘swim’:
HIW ᵍʟɔɣ (met.); LTG ɣɔ**r**; LHI ɣɒ**j**; VLW ɣɛ**j**; MTP ɣɛ**j**; VRS ɣœ**r**; MSN ɣɔ**r**; MTA ɣa**r**; NUM ɣa**r**

(#33) POC *ma-Raqan > PNCV *mara?a ‘light in weight’:
LYP mam**j**a; VLW mam**j**a; MTP mɔm**j**a

(#34) POC *[ma-]poRak > PNCV *ma-vora ‘broken, split’:
HIW wɔᵍʟə ‘break’; LHI mwɔ**j**wɒ**j**; LYP mɔwɔ**j**wɔ**j**; VLW mɔwɔ**j**wɔ**j**; MTP wɔ**j** ‘asunder’, mɔwɔ**j**wɔ**j**; LMG mɔwɔ**r**wɔ**r**; VRA mɔwɔ**r**wɔrɔ; VRS wɔ**r** ‘break’, mɔwɔ**r**wɔ**r**; MSN mɔwɔ**r**wɔ**r**; MTA mawo**r**a; NUM mɔmmɔɣɔ**r**; DRG mɔmɣɔ**r**; OLR mʊɣmʊɣ**r**a; MRL (ŋʷɔŋŋʷɔw)

(#35) PEOC *mʷeRa > PNCV *mʷera ‘child’:
HIW ŋʷaᵍʟə ‘baby’; LTG ŋʷe**r**ə ‘baby’; LHI sus\ŋʷæ**j**; LYP n-ʧi\n͡ŋ͡mʷɛ**j** ‘someone’; VLW nɛt\ŋmʷɛ**j**; MTP nɪt\n͡ŋ͡mʷɛ**j**; VRA ŋ͡mʷɛ**r**ŋmʷɛrɛ; VRS ŋ͡mʷi**r**ŋmʷia**r**; MSN ŋ͡mʷɛ**r**ŋmʷɛ**r**; MTA rɛrɛ\ŋmʷe**r**a ‘children’; NUM ŋ͡mʷaŋŋ͡mʷa**r**i; DRG ŋ͡mʷɛ**r**ŋ͡mʷɛ**r**; KRO rar\ŋ͡mʷɛ**r** ‘those two’; LKN (miːniˑ); MRL nʊ-lʊ\ŋʷɪ**r**

(#36) POC *ŋiRac > PEOC *ŋiRa ‘*Pemphis acidula*’:
HIW ŋeᵍʟə; LTG ŋɛ**r**ə; MTP nɪ-ŋɪ**j**; VRA ŋi**r**ɪ; VRS ŋi**r**; MSN ŋɪ**r**; MTA ŋi**r**a

(#37) PEOC *ŋoRo > PNCV *ŋoro ‘snore’:
MTP ŋɔ**j**ŋɔ**j**; LMG ŋør**r**ør; VRA ŋʊ**r**ŋʊ**r**; VRS ŋʊ**r**ŋʊ**r**; MTA ŋo**r**a; NUM ŋɔ**r** ‘sleep’; DRG ŋɔ**r** ‘sleep’; KRO ŋɔ**r** ‘sleep’; OLR ŋɔ**j** ‘sleep’; LKN ŋɔː ‘sleep’; MRL ŋɔ**r** ‘sleep’

(#38) PEOC *puRa > PNCV *bura ‘elephantiasis’:
LHI n-pɔ**j**; MTP -ᵐbʊ**j**, MTP ᵐbʊ**j**ᵐbʊ**j** ‘fat’; VRA ᵐbu**r**ʊ; MTA pu**r**a; MRL ᵐbɔ**r**

(#39) POC *qaRa(r) > PNCV *ara ‘fence’:
HIW (pa**j**a); LTG (pialə); LHI n-a**j**; LYP n-ɣ\a**j**; VLW nɛ-ɣɛ\a**j**; MTP nɛ-ɣɛ\a**j**; LMG n-ɒ**r**; VRS a**r**; MSN ɛrɛ; MTA ɣɛ\a**r**a; NUM ɣɛ\a**r**; DRG a**r**; LKN a**r**

(#40) POC *RapiRapi > PNCV *raviravi ‘afternoon, evening’:
HIW ᵍʟəβᵍʟɔβ; LTG rəβrɛβ; LHI ɟɪpɟæp; LYP ɟɪpɟɪp; VLW lɛm\ɟɛpɟɛp; MTP lɛm\ɟɪpɟɛp; LMG rɛβrɛβ; VRA rɛβrɛβ; VRS rɛβrɛβ; MSN rɛβrɛβ; MTA raβraβ; NUM rɛβrɛβ; DRG raːβrɪβ; KRO rɛβrɛaβ; OLR raβraβ; LKN ræβræβ; MRL rɛprɛp

(#41) PEOC *Rav[e,i] > PNCV *rave ‘pull, draw; angle (for fish)’:
HIW ᵍʟaβ; LTG raβ; LHI ɟɛp; LYP ɟɛp; VLW ɟɛp; MTP ɟap; LMG rɛβ; VRA rɪβ; VRS rɛβ; MSN rɛβ; MTA raβe; NUM rɛβ; DRG rɛβ; KRO rɛβ; OLR rɛβ; MRL rɪp

(#42) POC *Ropok ‘dash, fly’ > PNCV *rovo ‘run, flow, jump, fly’ > PTB *rowo ‘[VI] dash, rush out, jump, fly away; [DIREC] outwards, seawards’:
HIW ᵍʟow; LTG row; LHI ɟɔw; VLW ɟɔ; MTP ɟɔw; LMG rœw ~ rɔw; VRA rʊw; VRS rʊw; MSN rɔw; MTA rowo; NUM rɔw; DRG rɔw; KRO rɔw; OLR rɔw; LKN rɔw; MRL rɔw

(#43) POC *saRe > PNCV *sare ‘tear, rip’:
HIW sɪᵍʟ; LTG hi**r**; LHI hɛ**j**; LYP sa**j**; MTP ha**j**; VRS sɪ**r**; MSN sa**r**; MTA sa**r**e

(#44) PEOC *saRi ‘k.o. spear’ > PNCV *sari ‘[N] spear; [VT] spear, stab+’:
HIW sɔᵍʟ [VT]; LTG hɛ**r** [VT]; MTP n-ɪ\sɛ**j** [N]; VRA sɛ**r** [N]; VRS sɛ**r** [N]; MTA sa**r** [N]; DRG s**r**i/tʊk [N]; KRO te͡aβ\se͡a**r** [N]

(#45) POC *saRum ‘needle’ > PNCV *saru ‘sea urchin’:

HIW tət\sɔᵍʟ; LTG hiɛ**r**; MTP nɛ-hɛ**j**; MTA sa**r**

(#46) POC *sinaR > PNCV *ziŋa(ri) ‘shine, cast light’:
MTP hiŋ**j**ɛɣ ‘illuminate’; MTA siŋa**r**(aɣ) ‘illuminate, throw light upon’; NUM wi\sŋa**r** ‘moon, month’; DRG sŋa:**r** ‘moon, month’; LKN siŋa: ‘moon’

(#47) PSO *mʷaRaki > PNCV *m͡ʷara(ki) ‘*Chalcophaps indica*, k.o. dove’:
MTP na-ŋ͡mʷa**j**; LMG ŋ͡mʷɛ**r**ɛɣ; VRA ŋmʷa**r**aɣ; VRS ŋ͡mʷɛ**r**ɛɣ; MTA ŋ͡mʷa**r**a; DRG wɔ/ŋ͡mʷ**r**a:ɣ; LKN ŋ͡mʷæ**r**æɣ

(#48) PEOC *ʔu(c,z)uRi > PNCV *usu(ri) ‘follow (along)’ > PTB *(l)usuri ‘[v.] follow, [prep.] along’:
HIW ʉᵍʟ; LTG ʉ**r**i; LHI luhu**j**; LYP lilsi**j**; VLW lʊlhi**j**; MTP lʊlhi**j**; LMG lilsi**r**; VRA lisi**r**; VRS sü**r**; MSN su**r**; MTA usu**r**; NUM usu**r**; DRG sɪ**r**ɪ/ɣ; KRO su**r**; OLR tu\su**j**; LKN su:su:; MRL (ɣʉ)sʉ**r**

(#49) PEOC *ʔu(c,z)uRi > PNCV *usu(ri) ‘follow (along)’ > PTB *suri ‘[prep.] because of (s.t.); [conj.] because’:
HIW ʉᵍʟ; LTG ʉ**r**; LYP si**j** ‘dative’; VLW hi**j** ‘dative’; MTP hi**j** ‘dative’; LMG si**r**(ɛ); VRA si**r**(ɪn); VRS sü**r**; MSN su**r**; MTA su**r**; NUM su**r**; DRG su**r**; KRO su**r**; OLR su**j**; MRL sʉ**r**

(#50) PEOC *p(ʷ)aRaRa > PNCV *vara ‘handle’:
LYP n-βa**j**ɔ-; LMG n-βɒ**r**ɒ-; VRA βa**r**a-

(#51) POC *toRas > PNCV *tora ‘hardwood tree, esp. *Intsia bijuga*’:
MTP n-tɔ**j** ‘*Decaspermum neo-ebudicum*’; VRS tɔ**r**; MTA to**r**a; LKN tɔ:; MRL tɔ**r** ‘*Casuarina*’

(#52) POC *puRa ‘(water) bubble up’ > PNCV *vura ‘full’;[41]
HIW wuᵍʟə; LTG wu**r**ə; LHI wwɔ**j**; LYP ʊ**j**; VLW ʊ**j**; MTP ʊ**j**; VRS wʊ**r**; MTA u**r**a; DRG ββʊ**r**; OLR βʊβʊ**j**; LKN βʊβʊ:; MRL ffʊ**r**

(#53) POC *puRuk > PNCV *vuru ‘cough’:
MTP wu**j**; MTA βu**r**u

(#54) POC *wakaR > PNCV *kawa[ri] ‘root’:
LTG ɣə**r**əh; VLW n-ɣɪ**j**i; MTP nʊ-ɣʊ**j**i; LMG n-ɣœ**r**; VRA ɣʊ**r**i; VRS ɣɛ**r**i; MTA ɣa**r**i; DRG ɣa**r**i; KRO ɣɛa**r**; OLR ɣa**j**; LKN ɣæ:; MRL ɣɔɛ**r**

The following sets are not mentioned by Clark, but Lynch (n.d.) provides forms for Malakula. All these etyma show retention of *R in all NCV languages documented.

(#55) POC *biRapa > PNCV *birava ‘surgeonfish, *Acanthurus lineatus*’:
NUM ᵐbi**r**aβ; DRG ᵐb**r**aβ; KRO ᵐbi**r**aβ; LKN pi**r**aβ. [*R retained in Malakula: e.g., Nese ne\b̈i**r**av.]

(#56) POC *magaRut > PNCV *magaru ‘flying-fish, *Exocetus*’:
HIW ŋ͡ʷɔkɔᵍʟ; LTG ŋ͡ʷək**r**; LYP n-ŋ͡mʷɛkɛ**j**; VLW n-ŋ͡mʷɛᵑgɛ**j**; MTP na-ŋ͡mʷkɛ**j**; VRS ŋmʷœkœ**r**; MTA maka**r**u; DRG ŋmʷka:**r**; KRO ŋ͡mʷɛkɛa**r**; OLR ŋmʷɔka**j**; MRL ŋʷɔkɔ**r**; … Raga maga**r**u; Uripiv make**r**; Namakir maᵑga**r**/atah

(#57) POC *poRos ‘squeeze, wring’ > PNCV *woro ‘squeeze liquid from (grated coconut)’:
HIW woᵍʟ; MTP wɔ**j**; LMG wœ**r**; VRA wʊ**r**; VRS wʊ**r**; MTA wo**r**o; … cf. Lolovoli wo**r**o. [*R retained everywhere in Malakula: e.g.,Unua vu**r**uh…; but pockets of loss, or doublet showing loss.]

(#58) POC *(ma)Raŋo ‘withered’ > PNCV *(ma)raŋ[o,u] ‘(plant) dry, withered’:
HIW ᵍʟɛŋʷ ‘(yam) dry’; LTG re**r**ŋ ‘(yam) dry’; LHI mɪ**j**æŋ; VLW mɛ**j**ɛŋ; MTP mɛ**j**ɛŋ; LMG ma**r**aŋ; VRS mɛ**r**ɛŋ; MSN ma**r**aŋ. [*R retained everywhere in Malakula: e.g., Unua m**r**aŋ.]

(#59) PEOC *vaRa- ‘lungs’ > PNCV *vara- ‘chest; liver’:

41. The semantic link between PEOC *vuRa ‘(water) spring’ and PNCV *vura ‘full’ is proposed by Clark (2009).

HIW βəᵍʟɔ-, βaᵍʟə; LTG βəre-; MTP na-pjɛ-, na-βaj; LMG βœrœ-, βɒr; MSN βɔrɔ-; MTA βara-; NUM βara-; DRG βra:-; LKN βa:rɪ-; MRL nɵ-βɐrɔ-. [*R retained in Malakula: e.g., Port Sandwich na\var.]

APPENDIX 2. SOME LEXICAL ITEMS REQUIRING MORE DISCUSSION

A handful of items need to be examined in more detail, as they seem, at least at first sight, to show a more complex or patchy distribution of *R reflexes. These items are: *tabiRa 'bowl'; *buRaka '*Cyrtosperma* taro'; *piRu 'fan palm'.

1. *tabiRa 'wooden bowl'

Ross, Pawley, and Osmond (1998:72) reconstruct POC *tabiRa 'wooden bowl' after Geraghty (1990). As far as North Vanuatu languages are concerned, this root is apparently the source of two, or perhaps three, different words, which reflect *R in three different ways.

In appendix 1, I cited (#18) PNCV *tabia 'wooden dish where pudding is pounded', as an example where *R is lost throughout Vanuatu. A distinct word, which I reconstruct as PNCV *tabera 'a round, fine-woven basket used as a dish', is both close and different in meaning. It is probably a reflex of *tabiRa, with an irregular change of vowels. This *tabera retains *R everywhere in northern Vanuatu (at least for the cognates I have):

(#60) POC *tabiRa 'wooden bowl' > PNCV *tabera 'a round, fine-woven basket, used as a dish':
MTP na-tᵐbɛ**j**; VRS tiᵐbi͡a**r**; MSN tɛpɛ**r**; MTA tape**r**a; NUM tɛᵐbɛ**r**; … Raga taᵐbe**r**a 'plate'

Finally, a third word could be PTB *tabela 'triangular, coarse woven basket for rubbish'. Both formal and semantic similarity with (#60) make it a good candidate for being its etymological doublet:

(#61) POC *tabiRa 'wooden bowl' > (?) PTB *tabela 'a triangular, coarse woven basket for rubbish':
HIW pejə; LTG təpɛlə; OLR tɛpɛl; LKN sapɛl

If the etymology is correct, then this form *tabela would be the only example in the whole northern Vanuatu area where *R is reflected as *l in more than one language.[42] While this is a common reflex in various parts of the Pacific, it is an irregular one in this part of Vanuatu. The possibility of lexical borrowing cannot be eliminated (but from which donor?).

Even though these three forms *tabia, *tabeRa, and *tabela may have the same origin, they form distinct lexical sets (see the discussion in 2.2.3). Each set either loses *R or retains *R in a consistent way.

2. *buRaka 'swamp taro'

A lot has been written on the name of the swamp taro (*Cyrtosperma merkusii*) in Oceanic languages. Geraghty (1990:57, 2004:88) reconstructed PEOC *buRaka for this; Ross (1996) proposed POC *bulaka, but later came to more agnostic views on the nature of the liquid (Ross 2008:270). Kikusawa (2003) challenged the possibility of reconstructing any early etymon altogether, arguing that *Cyrtosperma* taro was introduced from Micronesia in more recent times into various Pacific regions. Her suggestion is that most (all?) non-Micronesian terms ultimately were borrowings from a Micronesian form (Proto–Chuukic-Ponapeic) *pʷulaka.

Even northern Vanuatu languages offer a complex picture. On the one hand, Geraghty (1990:57) and Clark (2009:86) reconstruct, for PNCV, an etymon *buaga 'taro, taro swamp'. The absence of any consonant between *u and *a is not compatible with an *l ety-

42. The glide /j/ is the regular reflex of *l in Hiw: POC *jalan 'road' > *ajə*; *alap 'take' > *əjə*. The loss of the first syllable (*tabela > *pejə*) is irregular, though.

mon, and thus makes *buRaka the most likely protoform (but note the irregular change *k > *g). This etymon is also reflected in the Torres-Banks area:[43]

(#62) POc *buRaka 'swamp taro' > PNCV *buaga 'swamp taro; taro swamp':
LHI n-pɪak 'swamp'; LYP n-paŋ 'taro swamp, boggy ground'; LMG n-puak 'mud'; VRS ᵐbʊak 'taro garden planted in mud'; MSN puak 'mud'; MTA puaka 'boggy ground, mud'; MRL nu-ᵐbuak 'mud'

If one combines this cognate set with other Vanuatu data (like, for example, Uripiv *bbuak* and Namakir *buag*), one may conclude that *R in *buRaka is lost everywhere in the archipelago.

This being said, it appears that a handful of northern languages have kept a trace of this *R in a doublet form that designates the swamp taro itself. *Cyrtosperma* taro seems to be only cultivated in the island of Ureparapara. In other North Vanuatu islands, it is little known. Hiw speakers have lost a name for it; as for Lo-Toga and Mwotlap, only a few elder speakers were able to remember it; I never heard any name in islands further south, and it is not reported in any of my other sources.

(#63) POc *buRaka > PTB *buraɣa '*Cyrtosperma* taro':
LTG pə**l**aɣə; LHI n-pɪ**j**a; LYP n-pa**j**a; MTP nɪ-m**j**aɣ

Interestingly for our purpose, these four forms show retention of *R. Only the /l/ in the Lo-Toga form is an irregular reflex; this may be due indeed to a borrowing from a language having /l/, perhaps from *pʷulaka. Nevertheless, the last three forms are perfectly regular reflexes of a form *buraɣa with a rhotic,[44] and point unambiguously toward an etymon *buRaka rather than *pʷulaka. In sum, the evidence from this small set of languages remains ambiguous as to whether the name for *Cyrtosperma* taro is a loanword (as suggested by the Lo-Toga form) or indigenous (as suggested by the three other languages).

Incidentally, the two languages of Ureparapara have preserved both etyma, one showing loss of *R (*buaga > LHI *n-pɪak*; LYP *n-paŋ*), the other showing its retention (*buRaka > LHI *n-pɪ**j**a*; LYP *n-pa**j**a*). This is evidence that the two cognate sets (#62) and (#63) must be considered distinct words in synchrony, with distinct meanings, as was the case with etymological doublets more generally (see 2.2.3). They therefore do not constitute any counterexample to the tendencies observed elsewhere in this paper, in terms of the distribution of *R reflexes across the territory. Ideally, more data should be added from other Southern Oceanic languages.

3. *piRu 'fan palm, *Licuala* spp.'

Proto-Oceanic had *piRu 'fan palm, *Licuala* spp.' The modern names of the plant in Torres-Banks languages appear in (#64):

(#64) HIW (ti) tɯw; LTG (hɔ) ʈɯβ; LYP n-(ⁿdø) ⁿdɯp; VLW n-jɛplɔɣ; MTP nɪ-jɪⁿdɪp; LMG n-tøɣɪβ; VRA ⁿdʊuβ; VRS ⁿdiβiβi; MTA βiloɣ; NUM ⁿdɔβlɛ; DRG ⁿda:βlʊ; KRO ⁿdɛ͡aβlʊ; OLR ʧaβi; LKN ʧiβi; MRL nʊ-ⁿdʊβlʊɣ

Despite their variety, these modern forms may, in fact, all be cognate, via three distinct intermediate protoforms.

First, all these forms (except Mota) incorporate an initial element that reflects POc *raun 'leaf', via an irregular form PTB *ⁿrau- (> *ⁿdau- in most languages). Three languages reflect a protoform *dau-βiu 'leaf of *Licuala*', in which the second element is obviously *piRu with a lost *R. Regular processes of unstressed vowel deletion (François 2005) explain why *-βiu is reflected as /-βi/ in these languages:

(#65) POc *rau(n) + piRu > *dauβiu 'fan palm, Licuala spp.':
VRS ⁿdiβiβi; OLR ʧaβi; LKN ʧiβi

43. The Vurës form is from Hyslop-Malau (n.d.), Mota is from Codrington and Palmer (1896), Nume and Mwesen are from Tryon (1976:340), and other languages are from my fieldnotes.

44. Both Lehali and Löyöp regularly lose /ɣ/ (< *k) in syllable-final position: see the reflexes of (#8) *baReko.

In six other languages, the four-syllable word *dauβíu, which had then become unanalyzable, was further shortened—irregularly—into a three-syllable form *daúβi, with reassignment of stress to the preceding syllable. Later, the regular process of unstressed vowel deletion reduced this *daúβi to two syllables (VRA *ⁿdʊuβ*, LMG *n-tøɣɪβ*), or even to one syllable (LTG *ʈɯβ*, MTP *-ⁿdɪp*):

(#65′) *dauβíu > *daúβi ‘fan palm, *Licuala* spp.’:

HIW (ti) tɯw; LTG (hɔ) ʈɯβ; LYP n-(ⁿdø) ⁿdɯp; MTP nɪ-jɪ\ⁿdɪp; LMG n-tøɣɪβ; VRA ⁿdʊuβ

In languages where *daúβi became a monosyllable, the latter is typically reinforced by the ordinary noun for ‘leaf’, itself a reflex of POC *raun> PTB *ⁿrau-. This is how Löyöp, for example, incorporates the etymon ‘leaf’ twice: *n-ⁿdø ⁿdɯp* < *na ⁿdáu ⁿdaú-βi.

Finally, a third set of modern forms appear to reflect an etymon *(ⁿrau-)βiloɣi:

(#66) *(rau-)βiloɣi ‘fan palm, *Licuala* spp.’:

VLW n-jɛ\ploɣ; MTA βiloɣ; NUM ⁿdɔ\βlɛ; DRG ⁿda:\βlʊ; KRO ⁿd͡ɛa\βlʊ; MRL ⁿdʊ\βlʊɣ

Geraghty (1990:86), and after him Ross (2008:222), suggest the latter forms could all be direct reflexes of *piRu, via *R > /l/; Ross also hypothesizes a final *q in the POC etymon—*piRu(q)—assuming this *q is the origin of /ɣ/ in *βiloɣi. Should Ross’s hypothesis be correct, then *piRu(q) would be (a) the only etymon in the Torres-Banks area where *q > /ɣ/ (everywhere else, *q > Ø); (b) almost the only word where *R > /l/ (but see [#60′] above); and (c) the only etymon for which *R is retained in a patchy way, rather than in a binary (south vs. north) isogloss. If one adds the mismatch of vowels, *piRu(q) > *βiloɣi, as a further concern, there are several reasons to treat Ross’s hypothesis with caution.

In sum, the only unambiguous cases where POC *piRu is reflected in north Vanuatu are those where *piRu shows loss of *R (> *βiu) and further irregular reduction to a monosyllabic formative *βi—cf. (#65′). It is not impossible that the first syllable *βi- in *βiloɣi may be the very same element, and therefore ultimately reflect *piRu. However, it seems that the final formative (< *-loki?) has a distinct and unknown origin, and cannot be straightforwardly attributed to *piRu. Geraghty (1990:86) mentions the two distinct forms of Gela, *vilu* and *vilo*, referring to two distinct kinds of fan palms; it is possible that these two forms point to two distinct etyma.

As far as the present study is concerned, this conclusion means that POC *piRu ‘fan palm’ should be included in the list of etyma in which *R was lost everywhere in northern Vanuatu languages (2.3.1).

APPENDIX 3. *R-LOSS IN VANIKORO LANGUAGES

Section 3.3 probed the southern outreach of the spread of *R-loss, and showed that the phenomenon essentially affected the same lexical items across Vanuatu and New Caledonia, following a consistent north-to-south cline. I have said little on the northern boundary of the phenomenon.

The first sections of this paper treated North Vanuatu as though it were the northernmost frontier for *R-loss. However, *R is also lost sporadically in regions further north in the Pacific: Southeast Solomonic, Micronesian, and Temotu (not to mention Central Pacific). The question is whether these are cases of parallel innovation, or if one can demonstrate continuity between these cases and the vast area covered in the present paper, running from New Caledonia to the Torres Islands. The cornerstone for this question, to put it simply, will be whether *R was lost in the same lexical items or not (see 4.4).

As far as Southeast Solomonic and Micronesian are concerned, this issue should be postponed to future studies; some answers may be gathered from the rich data in Geraghty (1990). Yet I would like to make some brief observations about one area that Geraghty did not cover in his initial study, namely Temotu. Ross and Næss (2007) identified the Temotu subgroup,

consisting of two branches, Reefs-Santa Cruz and Utupua-Vanikoro; they suggested it may have been a first-order subgroup of Oceanic (see figure 1). If this proposal is correct, and if one follows strictly the family-tree model, then one should not expect to find any shared innovation between Temotu languages and Southern Oceanic languages. However, since my conclusions in section 4 involve language diffusion anyway, it might well be the case that the nearby populations of Temotu and of the Torres-Banks had contact with each other at a time when their dialects of Oceanic were still mutually intelligible. Such ancient contact has never been demonstrated so far on a linguistic basis, but the observation of *R-loss might provide the key to this historical question. However, as the data below will show, the hypothesis of a possible connection does not seem conclusive, at least not with respect to *R.

I will illustrate Temotu with data from the three languages of Vanikoro on which I collected data in 2005 (François 2009): Teanu (TEA), also known as Buma; Lovono (LVN), also known as Vano; and Tanema (TNM). A survey of other Temotu languages must await future studies. Vanikoro is the island closest to the Torres Islands, and if some contact ever occurred between the two areas, it would probably be reflected in its languages.

The modern languages of Vanikoro have sometimes lost *R, but more often reflect *R with a segment. Just as for Vanuatu languages (see 2.2.1), the regular nonzero reflexes of *R are the same as *r, namely:

- word-internally, *r ~ *R > *l*: *meRaq 'red' > TEA ***mo*****l****oe**;
- when *r ~ *R was word-final, then it is reflected as an apical trill /r/, followed by a paragogic (nonetymological) vowel V:[45] *waiR 'water' > TEA ***e*****r****o**, LVN ***wi*****r****e**.

The latter case (*R followed by a paragogic vowel) includes a number of POC etyma where *R was word-final:

(#67) Etyma retaining word-final *R in Vanikoro:

*niuR 'coconut'	> TEA luro
*qauR 'bamboo'	> TEA okoro, TNM okaure
*waiR 'water'	> TEA ero, LVN wire
*toŋoR 'mangrove'	> LVN leŋore (but TEA toŋ^{m}be)
*qapaliR '*Acanthurus*'	> TEA wa\viliro
*bunaR 'Platax sp.'	> TEA ^{m}bunero
but	
*roŋoR 'hear, feel'	> TEA, LVN leŋi, TNM laŋi
*maturuR 'sleep'	> TEA mokoiu, TNM matou

This retention of word-final consonants distinguishes Temotu languages from their southern neighbors (see 2.2.2), and can therefore not be taken as a test in our comparative survey. However, the change *R > /r/ in Vanikoro languages is also found in a number of words that lost their original word-final vowel (often a high vowel) at some point of their history. An *R that became final because of this vowel loss behaved like the ones above when the merger with *r took place, changing into /r/ and adding a paragogic vowel: *paRi 'stingray' > *vaR# > *var > *varə > TEA *voro*. Thus compare the behavior of *R word-medially and (quasi-)word-finally in *biRibiRi 'Hernandia' > *biRibiR# > *bilibir > *bilibirə > TEA *^{m}bili^{m}biro*.

Vanikoro languages often retain *R in the same words as those of Vanuatu, whether this is true of the whole NCV group (#68), or at least of its northern area (#69):

(#68) Etyma retaining *R both across NCV and in Vanikoro:

*biRibiRi *'Hernandia'*	> TEA ^{m}bili^{m}biro
*yaRu '*Casuarina*'	> TEA iero
*qaRa 'fence'	> TEA aero, LVN ere, TNM eira
*Rabia 'starch'	> TEA leve 'arrowroot, *Tacca* sp.'
*mawiRi 'left hand'	> TEA mouro

45. This nonetymological, paragogic vowel is often /o/ in Teanu, /e/ in Lovono, /a/ in Tanema. I propose to reconstruct it as a schwa in Proto-Vanikoro (François 2009:108).

*wakaR 'root' > TEA kara, TNM ke

(#69) Etyma retaining *R north of NCV, and in Vanikoro:
*bakuRa '*Calophyllum*' > TEA ᵐburo
*buRaka '*Cyrtosperma*' > TEA vivilo, TNM peire
*goRu '(plant) dry' > TEA kokoro, LVN ve\ⁿgore, TNM va\ⁿgora
*[ka]ŋaRi '*Canarium*' > TEA vo\ŋoro, LVN ve\ŋere, TNM vi\ŋara
*kaRi 'shellfish' > TEA aero
*paRi 'stingray' > TEA voro
*paRu 'hibiscus' > TEA voro, LVN vel\vara
*puRa 'full' > TNM pura
*puRe 'k.o. creeper' > TEA ule
*quRis '*Spondias*' > TEA iliro
*suRi 'bone' > TEA, LVN ⁿdie, TNM ⁿdele (?)

However, these are all cases of SHARED RETENTION, so little can be said about them.

In several cases, Vanikoro retains *R where Vanuatu languages have lost it:

(#70) Etyma losing *R across NCV, retaining it in Vanikoro:
*baReko 'breadfruit' > TEA ᵐbale, TNM ᵐbaloe
*piRu 'fan palm' > TEA viro, LVN vire
*tapuRiq 'conch' > TEA teveliko, TNM lavlika
*meRaq 'red' > TEA moloe
*Rabia 'starch' > TEA ve\leᵐbie 'starch'
*jiRi '*Cordyline*' > TEA me\ⁿdiro

This could be seen as counterevidence to the hypothesis that Vanikoro might have been affected by the wave of *R-loss spreading across the whole Vanuatu archipelago. However, technically, this is not a solid argument. What (#70) shows is simply some lexical items whose *R-loss isogloss runs between Vanikoro and Vanuatu, in a way similar to those drawn on map 3. For these words, the northward spread of the *R-less variant has simply not reached the Temotu area.

More instructive are those cases where *R is lost both in NCV and in Vanikoro. However, I have identified only two such cases:

(#71) Etyma losing *R both across NCV and in Vanikoro:
*piRaq '*Alocasia* sp.' > TEA vioe, LVN piwe, TNM va
*Rumʷaq 'house' > TEA, LVN moe ~ mʷoe

To these two words, one may add the interesting case of ⟨56⟩ *vaRovaRo, which has lost its *R both in Vanikoro and in northern Torres-Banks, but not further south:

(#72) Etyma losing *R north of NCV, and in Vanikoro:
⟨56⟩ *vaRovaRo '*Neisosperma* sp.' > TEA vovo

Should these three examples be interpreted as cases of diffusion of *R-less forms from the south, this would entail the existence of linguistic relations between the Oceanic populations of Temotu/Vanikoro and those of North Vanuatu—something that has never been shown to date. This is not implausible, for various reasons that are geographic and historical, as well as linguistic, but I will not address them here. This being said, the small number of etyma in which the loss of *R is found in the two regions alike does not provide sufficiently solid evidence toward such a conclusion. One cannot rule out, for these three words, the possibility of parallel innovation.

A further argument tends to weigh in the direction of parallel innovation. This is the possibility, illustrated in (#73), for *R to be retained in northern Vanuatu, yet lost in Vanikoro:

(#73) Etyma retaining *R north of NCV, but losing it in Vanikoro:
⟨29⟩ *taRaq 'cut, chop' > TEA toe, LVN to, TNM teo
⟨35⟩ *kiRe 'pandanus' > TEA,TNM kie

⟨23⟩ *kaRat ‘bite’	> TEA, LVN, TNM ke[46]
⟨#26⟩ *boRe ‘dream’	> TEA mo^{m}bo, LVN me^{m}bu
⟨#37⟩ *ŋoR[a,o] ‘snore’	> TEA ŋoŋo

In sum, the hypothesis that *R-loss in specific items may have spread from northern Vanuatu across to Vanikoro is supported by three words, but challenged by five.

Consequently, if *R was sporadically lost in Vanikoro anyway, it may not be convincingly demonstrated that this formed part of the broader dialectological phenomenon of diffusion that encompassed the whole area of Southern Oceanic, from New Caledonia to northern Vanuatu. Based on current evidence, it remains equally likely that the loss of *R in Temotu languages was due to parallel innovation, with only a few coincidental cases, (#71) and (#72), where *R was lost in the same words in both areas. In sum, based on current evidence, the safest conclusion about the spread of *R-loss in specific individual words is probably not to include Temotu languages, as they might as well have gone through their own process of sporadic loss of *R.

APPENDIX 4. ABBREVIATIONS

The abbreviations for language names in the Torres and Banks islands were given in map 2. They are repeated below, together with other abbreviations.

DRG	Dorig	NUM	Nume
HIW	Hiw	OLR	Olrat
KRO	Koro	VLW	Volow
LHI	Lehali	VRA	Vera’a
LKN	Lakon	VRS	Vurës
LMG	Lemerig	POC	Proto-Oceanic
LTG	Lo-Toga	PROC	Proto-Remote Oceanic
LYP	Löyöp	PSO	Proto-Southern Oceanic
MRL	Mwerlap	PEOC	Proto-Eastern Oceanic
MTA	Mota	PNCV	Proto–North-Central Vanuatu
MTP	Mwotlap	PPN	Proto-Polynesian
MSN	Mwesen	PTB	Proto–Torres-Banks

46. Note, however, the form *kala* in Asumboa, a closely related language of Utupua (Tryon and Hackman 1983:426).

REFERENCES

Bedford, Stuart. 2003. The timing and nature of Lapita colonisation in Vanuatu: The haze begins to clear. In *Pacific Archaeology: Assessments and prospects*, ed. by Christophe Sand, 147–58. Nouméa: Les Cahiers de l'archéologie en Nouvelle-Calédonie, 15.

———. 2006. *Pieces of the Vanuatu puzzle: Archaeology of the north, south and centre*. Terra Australis 23. Canberra: Pandanus Press.

Bedford, Stuart, and Christophe Sand. 2007. Lapita and Western Pacific settlement: Progress, prospects and persistent problems. In *Oceanic explorations: Lapita and Western Pacific settlement*, ed. by Stuart Bedford, Christophe Sand, and Sean P. Connaughton, 1–15. Terra Australis 26. Canberra: ANU E-Press.

Bedford, Stuart, Christophe Sand, and Sean P. Connaughton, eds. 2007. *Oceanic explorations: Lapita and Western Pacific settlement*, Terra Australis 26. Canberra: ANU E-Press.

Bedford, Stuart, and Matthew Spriggs. 2008. Northern Vanuatu as a Pacific crossroads: The archaeology of discovery, interaction, and the emergence of the "ethnographic present." *Asian Perspectives* 47(1):95–120.

Blust, Robert. 2009. *The Austronesian languages*. Canberra: Pacific Linguistics.

Bril, Isabelle. 2000. *Dictionnaire nêlêmwa–nixumwak–français–anglais (Nouvelle-Calédonie)*. Langues et Cultures du Pacifique, 14. Paris: Peeters.

Bybee, Joan. 2002. Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language variation and change* 14:261–90.

Chambers, J. K., and Peter Trudgill. 1998. *Dialectology.* Cambridge: Cambridge University Press.

Charpentier, Jean-Michel. 1982. *Atlas linguistique du Sud-Malakula – Linguistic Atlas of South Malakula*, 2 vols. Paris: Société d'Etudes Linguistiques et Anthropologiques de France.

Charpentier, Jean-Michel, and Alexandre François. Forthcoming. *Atlas linguistique de Polynésie française – Linguistic atlas of French Polynesia*. Berlin: Mouton de Gruyter.

Clark, Ross. 1985. Languages of North and Central Vanuatu: Groups, chains, clusters and waves. In *Austronesian linguistics at the 15th Pacific Science Congress*, ed. by Andrew Pawley and Lois Carrington, 199–236. Canberra: Pacific Linguistics.

———. 2009. **Leo tuai: A comparative lexical study of North and Central Vanuatu languages*. Canberra: Pacific Linguistics.

Codrington, Robert H., and Jim Palmer. 1896. *A dictionary of the language of Mota, Sugarloaf Island, Banks Islands*. London: Society for Promoting Christian Knowledge.

Crowley, Terry. 2000. *An Erromangan (Sye) dictionary.* Canberra: Pacific Linguistics.

———. 2007. *The Avava language of central Malakula (Vanuatu)*, (ed. by John Lynch). Canberra: Pacific Linguistics.

Enfield, Nicholas. 2003. *Linguistic epidemiology: Semantics and grammar of language contact in mainland Southeast Asia*. London: Routledge-Curzon.

———. 2008. Transmission biases in linguistic epidemiology. *Journal of Language Contact* (Thema 2): 299–310.

Evans, Bethwyn, ed. 2009. *Discovering history through language: Papers in honour of Malcolm Ross*. Canberra: Pacific Linguistics.

François, Alexandre. 2001. Contraintes de structures et liberté dans l'organisation du discours. Une description du mwotlap, langue océanienne du Vanuatu. Doctoral diss., Université Paris-IV Sorbonne.

——. 2002. *Araki. A disappearing language of Vanuatu*. Canberra: Pacific Linguistics.

——. 2005. Unraveling the history of the vowels of seventeen northern Vanuatu languages. *Oceanic Linguistics* 44:443–504.
——. 2007. Noun articles in Torres and Banks languages: Conservation and innovation. In *Language description, history and development: Linguistic indulgence in memory of Terry Crowley*, ed. by Jeff Siegel, John Lynch, and Diana Eades, 313–26. New York: John Benjamins.
——. 2009. The languages of Vanikoro: Three lexicons and one grammar. In *Discovering history through language: Papers in honour of Malcolm Ross,* ed. by Bethwyn Evans, 103–26. Canberra: Pacific Linguistics.
——. 2010. Phonotactics and the prestopped velar lateral of Hiw: Resolving the ambiguity of a complex segment. *Phonology* 27(3):393–434.
——. Forthcoming. Social ecology and language history in the northern Vanuatu linkage: A tale of divergence and convergence. *Journal of Historical Linguistics*.
Galipaud, Jean-Christophe, and Bernard Vienne. 2005. *Chronologie du peuplement et réseaux d'échanges dans le nord du Vanuatu. Mission Santo 2005: Rapport préliminaire*. Nouméa: Institut de Recherche pour le Développement.
Garrett, Andrew. 2006. Convergence in the formation of Indo-European subgroups: Phylogeny and chronology. In *Phylogenetic methods and the prehistory of languages*, ed. by Peter Forster and Colin Renfrew, 139–51. Cambridge: McDonald Institute for Archaeological Research.
Geraghty, Paul A. 1978. Review article on D. T. Tryon, *New Hebrides languages: An internal classification*. *Kivung* 11(1):75–86.
———. 1983. *The history of the Fijian languages*. Oceanic Linguistics Special Publication 19. Honolulu: University of Hawai'i Press.
———. 1989. The reconstruction of Proto-Southern Oceanic. In *VICAL-1: Oceanic languages: Papers from the Fifth International Conference on Austronesian Linguistics*, ed. by Ray Harlow and Robin Hooper, 141–56. Auckland: Linguistic Society of New Zealand.
———. 1990. Proto-Eastern Oceanic *R and its reflexes. In *Pacific Island languages: Essays in honour of G. B. Milner*, edited by Jeremy H. C. S. Davidson, 51–93. London and Honolulu: School of Oriental and African Studies and University of Hawai'i Press.
———. 2004. Borrowed plants in Fiji and Polynesia: Some linguistic evidence. In *Borrowing: A Pacific perspective*, ed. by Jan Tent and Paul Geraghty, 65–98. Canberra: Pacific Linguistics.
Grace, George W. 1990. The "aberrant" (vs. "exemplary") Melanesian languages. In *Linguistic change and reconstruction methodology*, ed. by Philip Baldi, 155–73. Trends in Linguistics: Studies and monographs. Berlin: Mouton de Gruyter.
Gray, Russell D., A. J. Drummond, and Simon J. Greenhill. 2009. Language phylogenies reveal expansion pulses and pauses in Pacific settlement. *Science* 323 (5913):479–83.
Green, Roger C. 2003. The Lapita horizon and traditions—signature for one set of Oceanic migrations. In *Pacific Archaeology: Assessments and prospects*, ed. by Christophe Sand, 95–120. Nouméa: Les Cahiers de l'archéologie en Nouvelle-Calédonie 15.
Haudricourt, André-Georges, and Françoise Ozanne-Rivierre. 1982. *Dictionnaire thématique des langues de la région de Hienghène (Nouvelle-Calédonie)*. LACITO-documents. Paris: SELAF.
Heggarty, Paul, Warren Maguire, and April McMahon. 2010. Splits or waves? Trees or webs? How divergence measures and network analysis can unravel language histories. *Philosophical Transactions of the Royal Society B: Biological Sciences* 365:3829–43.
Henri, Agnès. 2010. *Éléments de description d'une langue mélanésienne du Vanuatu, le sunwadia*. Doctoral diss., Université Paris-IV Sorbonne.

Hock, Hans Henrich. 1991. *Principles of historical linguistics*. Trends in Linguistics: Studies and Monographs 34. Berlin: Mouton de Gruyter.
Hock, Hans Henrich, and Brian D. Joseph. 1996. *Language history, language change, and language relationship: An introduction to historical and comparative linguistics*. Trends in Linguistics: Studies and Monographs 93. Berlin: Mouton de Gruyter.
Hyslop, Catriona. n.d.a. Vera'a–English lexicon. Electronic files. Research Centre for Linguistic Typology.
———. n.d.b. Vurës–English lexicon. Electronic files. Research Centre for Linguistic Typology.
———. n.d.c. Lolovoli–English lexicon. Electronic files. Australian National University.
Jauncey, Dorothy. n.d. Tamambo-English lexicon. Electronic files. Australian National University.
Kikusawa, Ritsuko. 2003. Did Proto-Oceanians cultivate Cyrtosperma taro? *People and Culture in Oceania* 19:26–54.
Kirch, Patrick V. 1997. *The Lapita peoples: Ancestors of the Oceanic world*. Cambridge: Blackwell.
Krishnamurti, Bh. 1998. Regularity of sound change through lexical diffusion: A study of s > h > Ø in Gondi dialects. *Language Variation and Change* 10:193–220.
Labov, William. 1963. The social motivation of sound change. *Word* 19:273–309.
———. 1994. *Principles of linguistic change: Internal factors*. Oxford: Blackwell.
———. 2001. *Principles of linguistic change: Social factors*. Oxford: Blackwell.
Le Page, R. B., and Andrée Tabouret-Keller. 1985. *Acts of identity: Creole-based approaches to language and ethnicity*. Cambridge: Cambridge University Press.
Lynch, John. 2000a. Linguistic subgrouping in Vanuatu and New Caledonia. In *Proceedings of the Second International Conference on Oceanic Linguistics (SICOL),* vol. 2: *Historical and descriptive studies*, ed. by Bill Palmer and Paul Geraghty, 155–84. Canberra: Pacific Linguistics.
———. 2000b. Reconstructing Proto-Oceanic stress. *Oceanic Linguistics* 39:53–82.
———. 2000c. South Efate phonological history. *Oceanic Linguistics* 39:320–38.
———. 2001. *The linguistic history of Southern Vanuatu*. Canberra: Pacific Linguistics.
———. 2005. Final consonants in Remote Oceanic. *Oceanic Linguistics* 44:90–112.
———. 2008. Liquid vocalization and loss in Central Vanuatu. *Oceanic Linguistics* 47:294–315.
———. 2009a. Irregular sound change and the post-velars in some Malakula languages. In *Austronesian historical linguistics and culture history: A festschrift for Bob Blust*, ed. by Alexander Adelaar an d Andrew Pawley, 57–72. Canberra: Pacific Linguistics.
———. 2009b. At sixes and sevens: The development of numeral systems in Vanuatu and New Caledonia. In *Discovering history through language: Papers in honour of Malcolm Ross*, ed. by Bethwyn Evans, 391–411. Canberra: Pacific Linguistics.
———. n.d. Retention and loss of *R in some Malakula languages. Unpublished ms.
Lynch, John, and Amanda Brotchie. 2010. Vowel loss in Tirax and the history of the apicolabial shift. *Oceanic Linguistics* 49:369–88.
Lynch, John, Malcolm Ross, and Terry Crowley. 2002. *The Oceanic languages*. Richmond: Curzon Press.
Marck, Jeff. 2000. *Topics in Polynesian language and culture history.* Canberra: Pacific Linguistics.
Milke, Wilhelm. 1958. Zur inneren Gliederung und geschichtlichen Stellung der Ozeanisch-Austronesischen Sprachen. *Zeitschrift für Ethnologie* 83:58–62.
Moyse-Faurie, Claire, and Marie-Adèle Néchérö-Jorédié. 1986. *Dictionnaire xârâcùù–français (Nouvelle-Calédonie)*. Nouméa: Edipop.
Ozanne-Rivierre, Françoise. 1984. *Dictionnaire iaai–français (Ouvéa, Nouvelle-Calédonie)*. Langues et cultures du Pacifique, 6. Paris: Société d'Études Linguistiques et Anthropologiques de France.

———. 1992. The Proto-Oceanic consonantal system and the languages of New Caledonia. *Oceanic linguistics* 31:191–207.
———. 1995. Structural changes in the languages of northern New Caledonia. *Oceanic Linguistics* 34:45–72.
———. 1998. *Le nyelâyu de Balade (Nouvelle-Calédonie*). Langues et cultures du Pacifique, 12. Paris: Peeters.
Pawley, Andrew. 1972. On the internal relationships of Eastern Oceanic languages. In *Studies in Oceanic culture history*, ed. by R. C. Green and M. Kelly, 1–142. Pacific Anthropological Records, 13. Honolulu: Bernice P. Bishop Museum.
———. 1973. Some problems in Proto-Oceanic grammar. *Oceanic Linguistics* 12:103–88.
———. 1981. Melanesian diversity and Polynesian homogeneity: A unified explanation for language. In *Studies in Pacific languages and cultures in honour of Bruce Biggs*, ed. by Jim Hollyman and Andrew Pawley, 259–310. Auckland: Linguistic Society of New Zealand.
———. 1999. Chasing rainbows: Implications of the rapid dispersal of Austronesian languages for subgrouping and reconstruction. In *Selected papers from the Eighth International Conference on Austronesian Linguistics*, ed. by Elizabeth Zeitoun and Paul Jen-Kuei Li, 95–138. Taipei:Symposium Series of the Institute of Linguistics, Academia Sinica.
———. 2007a. Why do Polynesian island groups have one language and Melanesian island groups have many? Patterns of interaction and diversification in the Austronesian colonization of Remote Oceania. Paper read at "Migrations" conference, September 2007, Porquerolles (Var, France).
———. 2007b. The origins of Early Lapita culture: The testimony of historical linguistics. In *Oceanic explorations: Lapita and Western Pacific settlement*, ed. by Stuart Bedford, Christophe Sand, and Sean P. Connaughton, 17–49. Terra Australis 26. Canberra: ANU E-Press.
Pawley, Andrew, and Malcolm Ross. 1995. The prehistory of Oceanic languages: A current view. In *The Austronesians: Historical and comparative perspectives*, ed. by Peter S. Bellwood, James J. Fox, and Darrell Tryon, 39–80. Canberra: Australian National University.
Rivierre, Jean-Claude. 1983. *Dictionnaire paicî–français, suivi d'un lexique français–paicî*. Paris: Société d'Études Linguistiques et Anthropologiques de France.
———. 1994. *Dictionnaire cèmuhî–français: suivi d'un lexique français–cèmuhî*. Société d'Etudes Linguistiques et Anthropologiques de France. Paris, Louvain: Peeters.
Ross, Malcolm. 1988. *Proto-Oceanic and the Austronesian languages of western Melanesia*. Canberra: Pacific Linguistics.
———. 1996. Reconstructing food plant terms and associated terminologies in Proto Oceanic. In *Oceanic Studies: Proceedings of the First International Conference on Oceanic Linguistics*, ed. by John Lynch and Fa'afo Pat, 163–221. Canberra: Pacific Linguistics.
———. 1997. Social networks and kinds of speech-community events. In *Archaeology and language 1: Theoretical and methodological orientations*, ed. by Roger Blench and Matthew Spriggs, 209–61. London: Routledge.
———. 1998. Proto Oceanic phonology and morphology. In *The lexicon of Proto Oceanic*. vol. 1: *Material culture*, edited by Malcolm Ross, Andrew Pawley, and Meredith Osmond, 15–35. Canberra: Pacific Linguistics.
Ross, Malcolm, and Åshild Næss. 2007. An Oceanic origin for Äiwoo, the language of the Reef Islands? *Oceanic Linguistics* 46:456–98.
Ross, Malcolm, Andrew Pawley, and Meredith Osmond, eds. 1998. *The lexicon of Proto Oceanic,* vol. 1: *Material culture*. Canberra: Pacific Linguistics

———. 2003. *The lexicon of Proto Oceanic*, vol. 2: *The physical environment*. Canberra: Pacific Linguistics.
———. 2008. *The lexicon of Proto Oceanic*, vol. 3: *Plants*. Canberra: Pacific Linguistics.
———. Forthcoming. *The lexicon of Proto Oceanic*, vol. 4: *Animals*. Canberra: Pacific Linguistics.
Sand, Christophe, ed. 2003. *Pacific archaeology: Assessments and prospects*. Nouméa: Les Cahiers de l'archéologie en Nouvelle-Calédonie 15.
Schmidt, Hans. n.d. Nume–English lexicon. Electronic files. London: SOAS.
Schmidt, Johannes. 1872. *Die Verwandtschaftsverhältnisse der indogermanischen Sprachen.* Weimar: Hermann Böhlau.
Shing, Richard, Stuart Bedford, Matthew Spriggs, and Ralph Regenvanu. 2007. *Ripot long saed blong stadi long akioloji long not blong Vanuatu (2005-2007).* Port Vila: Vanuatu National Library.
Spriggs, Matthew. 1997. *The island Melanesians*. Oxford: Blackwell.
Thieberger, Nicholas. 2007. Dictionary of South Efate. Electronic publication. [Available from http://paradisec.org.au/SELexicon/index-english/main.htm].
Tryon, Darrell T. 1976. *New Hebrides languages: An internal classification*. Canberra: Pacific Linguistics
———. 1996. Dialect chaining and the use of geographical space. In *Arts of Vanuatu*, ed. by Joël Bonnemaison, Kirk Huffman, Christian Kaufmann, and Darrell Tryon, 170–81. Bathurst: Crawford House Press.
Tryon, Darrell T., and B. D. Hackman. 1983. *Solomon Islands languages: An internal classification*. Canberra: Pacific Linguistics.
van den Berg, René. 2006. Retained and introduced final consonants in Vitu. *Oceanic Linguistics* 45:43–52.
Wang, William, ed. 1977. *The lexicon in phonological change*. The Hague: Mouton.
Wheatley, J. I. 1992. *A guide to the common trees of Vanuatu, with lists of their traditional uses and ni-Vanuatu names*. Port Vila: Vanuatu Department of Forestry.
Wivell, Richard. 1981. Kairiru grammar. MA thesis, University of Auckland.

francois@vjf.cnrs.fr