

Lexical history in the Northwest Solomonian languages

Malcolm Ross

Malcolm.Ross@anu.edu.au

March 6, 2008

Draft only. Please do not cite without the author's agreement.

1 Lexical history in the Northwest Solomonian languages

1 Introduction

Sheppard, Walter and Roga (this volume) summarise archaeological evidence for the settlement history of the Western Solomons (Mono and Alu, Choiseul, the New Georgia group and Santa Isabel). They refer to a proposal based on circumstantial linguistic evidence that I advanced twenty years ago to the effect that there had been two waves of Oceanic settlement in the Western Solomons (Ross 1988:382–386). The proposal has not been widely accepted among Oceanist linguists, and this paper seeks to offer, among other things, more direct linguistic support for the two-wave proposal.

Andrew Pawley, in whose honour this volume is published, has a longstanding interest in the linguistic history of the Solomon Islands, and especially of the Southeast Solomonian (SES) languages.¹ I shall say almost nothing here about the SES group, but the history of their immediate neighbours in the Northwest Solomonian (NWS) group must provide at least one piece in the SES historical jigsaw.

2 Northwest Solomonian: languages and abbreviations

NWS languages fall into six groups:²

- (1) a. North Bougainville (NBv):³ *Nehan, Solos, Petats, Haku, Selau, Taiof, Teop*
- b. Banoni–Piva (BP): *Banoni, Piva*

¹ I owe a very considerable debt of gratitude to Andy. It was one of his papers (Pawley 1975) which inspired my first foray into historical linguistic research, emulating his title (Ross 1977), and he was one of those who encouraged me in my late entry into the field. His arrival at the Australian National University came shortly after my appointment there, and he has been a mentor and friend ever since.

² The languages listed here are those referred to in this paper. For a full listing see Ross (1988:217) or, with a different nomenclature, Tryon & Hackman (1983). Maps showing the locations of languages can be found in Ross (1988), Ross et al. (1998) or Ross et al. (2003).

³ Labelled 'Nehan/North Bougainville' in Ross (1988).

- c. Mono-Torau (MT): *Mono, Torau, Uruava* (extinct)
- d. Choiseul (Ch):
 - i. West Choiseul (WCh): *Vaghua, Varisi, Ririo*
 - ii. East Choiseul (ECh): *Ririo, Babatana, Sisingga*
- e. New Georgia (NGe):
 - i. West New Georgia (WANGe): *Simbo, Lungga, Ghanongga, Nduke*
 - ii. East New Georgia (ENGe): *Nduke, Roviana, Ughele, Kusaghe, Hoava, Marovo* (Mvo), *Vangunu*
- f. Santa Isabel (Is): *Kia, Kokota, Laghu, Kilokaka (Zazao), Blablanga, Ghove, Maringe* (Mge)

The placement of Ririo in both West and East Choiseul and of Nduke in both West and East New Georgia is not an error, but a mark of the fact that each of these languages provides a transition between the two parts of its respective group. Abbreviations for language names consist of the first three letters of the name except in the cases of Marovo (Mvo) and Maringe (Mge). Abbreviations of protolanguage names are formed by placing lower-case ‘p’ before the abbreviation for the group name, e.g. pNWS for Proto Northwest Solomonic.

The close relatedness of Mono and Torau appears surprising in the light of their present locations, but until less than two centuries ago Torau speakers were evidently located in the extreme southeast of Bougainville, just across the water from Mono (Terrell & Irwin 1972).

3 Northwest Solomonic: genealogy

The NWS group belongs to the larger Western Oceanic linkage (a linkage is a group of languages that has emerged from an earlier dialect network). Western Oceanic comprises those languages of the Oceanic subgroup of Austronesian that are located on the north coast of West Papua, in Papua New Guinea excluding the Admiralties, and in the Western Solomons (Ross 1988:386–389, Lynch et al. 2002:96).

Proto Oceanic must have been spoken in an area which included New Britain and New Ireland (Pawley 2003). Its speakers were associated with the early phase of the archaeologically recognisable Lapita culture around 1400 BC. They spread early to the Admiralties and around 1200 BC made their first southeastward push via Bougainville into the Solomons and northern Vanuatu, whence they moved eastward to settle Fiji, Polynesia and Micronesia and southward to settle the Vanuatu archipelago, the Loyalties and New Caledonia.⁴

Western Oceanic appears to have diversified out of a dialect network that was left behind in New Britain and New Ireland. However, before it diversified, certain innovations took place in the network that were not shared by the speakers in the first southeastward push (Ross 1988:382–383, Lynch et al. 2002:101). In the course of its diversification, Western Oceanic split into two or three separate networks, with a major division at the Willaumez Peninsula on the north coast of New Britain.⁵ The network to the east of the peninsula has been dubbed the Meso-Melanesian (MM) linkage (Ross 1988:257, 423 footnote 98).⁶ The linkage has a complex internal structure, in outline as follows:⁷

⁴ For details, see Green (1995), Kirch (1997, 1998), Lynch et al. (2002:Ch 5), Pawley and Ross (1993, 1995), Spriggs (1995, 1997).

⁵ I argued in Ross (1988:382) that this was the location of the pOc homeland. Pawley (2005) rightly points out that the evidence says nothing directly about pOc, only about Western Oceanic.

⁶ Regarding its characterisation as a linkage, see Lynch et al. (2002:101).

⁷ Adapted from Lynch et al. (2002:101), where I replaced the clumsy ‘South New Ireland/Northwest Solomonic linkage’

- (2) Meso-Melanesian linkage
Bali, Vitu
 Willaumez linkage
 New Ireland/Northwest Solomonic linkage
 Tungag/Nalik family
 Tabar linkage
 Madak linkage
Tomoip
 St George linkage
 South New Ireland languages
 Northwest Solomonic linkage

The NWS linkage, then, is a portion of the St George linkage and is coordinate within it with a number of South New Ireland languages. The latter appear to form seven groups, each coordinate with NWS, but the detailed history of these relationships is complicated (Ross 1988:258, 306–314, 1997). Although NWS is labelled a linkage because its languages evidently emerged out of a dialect network (§7.1), the network itself resulted from the dispersal of speakers of a single language, pNWS. We can be confident of this because NWS languages reflect certain innovations not found in other MM languages. They were (Ross 1988:218, 247–249)

- (3) a. pOc *w was lost.
 b. A vowel was added after a pOc final consonant. This vowel echoed the vowel before the final consonant. For example, pOc *boRok ‘pig’ > pNWS *boroyo.
 c. Following 3b, pOc word-final *-q became pNWS *-k (elsewhere *q became pNWS *y, merging with pOc *k).⁸
 d. The pOc first-person singular free pronoun *[i]au acquired an accreted *r-, becoming *r[i]au.

With regard to 3d, accreted *r- is also mostly reflected on the other first- and second-person pronouns in the NBv, BP, MT and Ch groups, but not in NGe or Is. This accretion was probably the outcome of major changes in clause-order that occurred in pNWS (Ross 1988:228–247), but their discussion lies beyond the scope of this paper.

Certain other NWS innovations relative to pOc had already occurred in MM but are relevant to the interpretation of some of the data given below. There were three mergers: pOc *r and *R merged as pMM *r, pOc *dr and *d as pMM *d, and pOc *s and *c as pMM *s. There were also two splits: pOc *k split into pMM *y and *k, pOc *p into pMM *v and *p.⁹ NWS innovations are illustrated with supporting data in Ross (1986).

One complex morphosyntactic innovation receives frequent mention in the literature: this is the adoption of what were once possessive noun phrase structures as verb phrase structures (Ross 1982, Palmer 2002, 2003). This may indeed have been a pNWS innovation, but precisely because it is syntactic, it may also have arisen through contact and is thus not a strong candidate for shared inheritance from the protolanguage, *pace* Ross (1988:249–251). Indeed, it is also reflected in certain SES languages, perhaps as a result of contact.

of Ross (1988:258)) with the ‘St George linkage’, after Cape St George (the southernmost tip of New Ireland) and the St George’s Channel (between southern New Ireland and the Gazelle Peninsula of New Britain). Language names are in italics.

⁸ In the orthography of Ross (1988) pNWS *y was shown as pNWS *q.

⁹ The conditioning of the splits of pOc *k into pMM *k and *y and of pOc *p into pMM *p and *v is not well understood.

Each of the six NWS groups except Banoni-Piva is characterised by certain phonological and morphosyntactic innovations relative to pNWS, but in comparison with the innovations characterising NWS as a whole, they are rather insignificant. They are:

- (4) a. North Bougainville: pNWS *ɣ is lost; pNWS *u became pNBv *i in certain lexical items; pNBv innovated two noun classes, marked by articles *a and *u; the article is repeated before an attributive adjective (Ross 1988:223, 252–253).
- b. Banoni–Piva: none (relationship is obvious by inspection).
- c. Mono-Torau: pNWS *ɣ is lost; SOV clause order, preposed possessor, postpositions (Ross 1988:223, 253–255).
- d. Choiseul: pNWS *s is lost in some pCh items, retained as pCh *s in others (Ross 1988:224).¹⁰
- e. New Georgia: pNWS *sava ‘what?’ replaced by pNGe *saqa (Ross 1988:224).
- f. Santa Isabel: pNWS *v became pIs *f; pNWS medial *-s- often became pIs *-h-; PWNS *tolu ‘three’, *visa ‘how many?’ and *vai ‘where’ became respectively pIs *tilo (expected **tolu), *n-iha (expected **na-fiha) and *hae (expected *fae) (Ross 1988:225).

There are also innovations which are common to the New Georgia and Santa Isabel groups, suggesting that they may have had a short period of exclusively shared history. However, this hypothesis now appears much weaker than it did in 1988. Proto NWS *rani ‘day’ became *rane (Ross 1988:224). Proto NWS non-first-person-singular free pronouns with accreted *r- do not occur, as noted in connection with 3d: whether this is a shared innovation is not clear. I claimed in Ross (1988:215, 240–247) that Roviana (ENGe) and Maringe (Is) both reflected an innovatory postverbal topic marker *si. Evidence has since become available that it is not reflected in other NGe or Is languages and probably should not be attributed to a shared protolanguage.

4 Papuan languages of northwest Island Melanesia

Crucial to the discussion in this paper is the fact that the Bismarck Archipelago (New Britain, New Ireland and the Admiralty Islands), Bougainville and the Western Solomons were occupied by Papuan speakers for millennia before the arrival of speakers of Austronesian languages. In this context, ‘Papuan’ simply means ‘not Austronesian’, as there is reasonably good evidence in the form of the surviving Papuan languages of the region that by the time Austronesian speakers arrived, the various groups of Papuan languages had long since lost any indicators of genealogical relationship (assuming that such a relationship once existed) (Ross 2001, 2005, Dunn et al. 2002, 2005, Terrill 2002).

Much of this region was settled by speakers of (what we may assume to have been) ancestral Papuan languages during the Pleistocene (Ross 2006). Soon after 19,000 BC, after the Last Glacial Maximum, there was a shift in New Britain from mobile foraging to foraging sedentism (Spriggs 1997:61–65). There are indications that animal and plant species were deliberately imported into New Ireland and Manus. Spriggs (1996, 1997:31–34, 61) interprets this as the beginning of what he calls wild-food production, i.e. the deliberate tending of the forest environment by selective weeding or clearing and by transplanting, without the permanent clearing of the forest which is entailed in agriculture. This situation appears to have subsisted until the introduction of agriculture by Austronesian speakers.

¹⁰ Ross (1988:224) interpreted reflexes of pNWS *s as pCh *j as a further split. In §7.5 below they are treated as borrowings into pWCh.

5 The two-wave proposal

The Solomon Islands are bisected linguistically by a line identified by Tryon & Hackman (1983), which I dubbed the ‘Tryon–Hackman line’ in 1988. It forms the boundary between the NWS and SES groups. The line raises a historical question. Why do NWS and SES reflect quite different innovations? If both were descendants of languages which arrived in the Solomons during the rapid southeastward push around 1200 BC, we would expect them to show a closer relationship with each other than they do. In particular, the pOc phoneme *R has a history of instability. In MM, including NWS, *R merged with *r. In SES *R merged with *l.¹¹

How does one account for this and other differences between NWS and SES? Broadly, there are two possible kinds of hypothesis. The first says that the earliest NWS and SES speakers were both descended from settlements established during the rapid southeastward expansion around 1200 BC, but that speakers of early Oceanic remained in intense enough contact with each other for a long enough period of time for the innovations which characterise MM to spread through the whole of the early MM dialect network, including pNWS. I find the required intensity of contact over an extended period difficult to believe in. Furthermore, even if I did believe in it, I would still have to explain the hiatus in contact that is reflected in the Tryon–Hackman line.

Because of these difficulties, I continue to prefer a two-wave hypothesis like that put forward in Ross (1988:382–386), which suggests (i) that pSES and the languages of Remote Oceania are outcomes of the expansion around 1200 BC; (ii) that the innovations that characterise MM occurred somewhat later, in a smallish area centred somewhere on or near the Willaumez Peninsula of New Britain, as this is clearly the area of greatest diversity in the MM linkage;¹² (iii) that speakers of early MM languages then expanded east and northeast into New Ireland; and (iv) that speakers of a South New Ireland language and their descendants moved southeastward into Buka, Bougainville and the Western Solomons. This language became pNWS. For socio-economic reasons which probably entailed symbiotic relationships with Papuan speakers (cf Dutton 1994), the NWS southeastward expansion stopped roughly at the furthest point of much earlier Papuan expansion. At some point after this, NWS speakers came into contact with SES speakers and the Tryon-Hackman line came into being.

The summary by Sheppard et al. (this volume) indicates that some form of two-wave hypothesis enjoys archaeological support, in that work in the NWS region has turned up no signs of early Lapita (i.e. first wave) settlement, despite the presence of such settlements further east. Instead, there is evidence of late Lapita settlement on Buka around 800 BC and in the New Georgia island group around 600 BC, which seems to correlate with the second-wave southeastward spread of MM (Sheppard & Walter 2006).

It follows from the two-wave hypothesis that there were already some Oceanic speakers in New Ireland and the NWS region when MM speakers arrived there. They were the descendants of speakers who had settled during the first expansion around 1200 BC, and I shall refer to their languages as ‘Old Oceanic’ languages. One might well ask why no Old Oceanic languages have survived in New Ireland or the NWS region. Archaeological evidence indicates that initial settlement was rapid and that ‘early Lapita communities were largely confined to smaller offshore islands in the Bismarck Archipelago that offered rich lagoon resources and land where they could plant their crops and live at a safe distance from mainland indigeneous [Papuan-speaking] populations.’ (Pawley 2005). By 1100 BC Old Oceanic had reached the Reefs and Santa Cruz Islands, located between the Solomons and Vanuatu archipelagoes (Green 2003). The MM spread was, one may infer, more gradual and supported by a

¹¹ In (at least) much of Remote Oceanic, most occurrences of *R were lost, but it merged with *r in some lexical items (Geraghty 1990).

¹² See Map 5.3 in Lynch et al. (2002:103).

growing population of speakers, and involved more co-operative relationships with Papuan speakers.¹³ It is a reasonable inference that when Old Oceanic and MM speakers came into contact, the latter were socially dominant and more populous, and their languages prevailed. As Pawley (2005) points out, however, replacement was not a large-scale process, as Old Oceanic speakers in most cases probably did not live on the larger islands.

The linguistic evidence offered for the two-wave proposal has largely been circumstantial. If, however, MM languages do represent a second wave such as I have described, we would at least expect to find loanwords from both Papuan and Old Oceanic languages, and perhaps more radical changes resulting from language shift like those found in Madak and Lamusong on New Ireland (Ross 1994). The rest of this paper is a preliminary investigation of the lexicon of languages of the NWS subgroup of MM. It shows that (i) the lexical retention rates of NWS are lower than those of non-NWS Oceanic languages; (ii) there are numerous apparent Papuan loans in NWS, but they cannot readily be sourced; (iii) there are NWS lexical items which appear to be Old Oceanic loans.

6 The retention rates of Northwest Solomonian languages

If the scenario above is roughly correct, we would expect higher retention of pOc lexicon in SES than in NWS, and this is indeed what we find. I demonstrate this difference below by quantifying the degree to which basic vocabulary items in NWS and SES languages reflect reconstructable pOc etyma. I also examine the diversity of NWS lexicons, although there is no simple way to quantify this.

To determine the relative lexical conservatism of NWS and SES languages I calculated retention rates relative to pOc for most NWS languages, for a sample of SES languages (Bugotu, Gela, Tolo, Lau, Kwaio, Kwara'ae, To'aba'ita, Sa'a and Santa Ana), and, for comparison's sake, for a few Oceanic languages outside the Solomons (Gedaged, Motu, Vitu, Tigak, Tabar, Lihir, Kandas, Mota and Bauan Fijian).

The procedure for calculating retention rates is modelled on that used by Blust (1981). He reconstructs Proto Malayo-Polynesian etyma for a modified version of the Swadesh 200-meaning list, calculating the percentage of these etyma reflected in each of the 55 languages in his database. Since all the 40 languages in my database are Oceanic, I instead used a baseline list of pOc etyma reflecting recent research, mainly Ross, Pawley and Osmond (1998, 2003) and, for free pronouns, Lynch et al. (2002:ch 4). Following Blust (2000), I allowed more than one etymon per meaning where there is no discernible difference in meaning between reconstructed etyma. I modified Blust's meaning list in various small ways¹⁴ and ran a trial with the resulting list of 199 meanings.

The trial revealed a number of weaknesses in the data set that a more constrained version of the meaning list could avoid. Space precludes a detailed account of these weaknesses, but they entailed ambiguity, polysemy and conceptual mismatch,¹⁵ and, as a result of these characteristics, there was either a plethora of reconstructions for a given item or insufficient or incorrect attestation of the required meanings in the wordlists. Meanings with these characteristics tend to display low retention in

¹³ This inferred difference in social relations between Old Oceanic and MM speakers receives some support from Wickler's (2001:241) interpretation of the archaeological sequence on Nehan and Buka islands in the North Bougainville area.

¹⁴ The list can be found at the Austronesian Basic Vocabulary Database website (<http://language.psy.auckland.ac.nz/austronesian/>). The website's authors have added numerals from 6 to 10 and 'a hundred'. I removed 'this', 'that', 'and', 'no', 'if', 'other' and 'all', as many Oceanic languages have more than one word corresponding to each and there is no obvious way of standardising one's choice across languages. In keeping with Oceanic lexical organisation 'we' was replaced by 'we (inclusive)' and 'we (exclusive)' and 'wife' by 'spouse', and excluded 'salt' as it is often not conceptually separable from 'sea (water)'.

¹⁵ Conceptual mismatch refers to cases where the English meaning elicits more than one Oceanic concept, or no Oceanic concept at all.

Table 1: Summary of retentions from Proto Oceanic for 106 meanings in 40 languages

	No. of meanings	No. of entries	Retention rate (%)	Reflexes of pOc items (%)
Whole database	102.4	112.7	46.8	49.7
North New Guinea				
<i>Gedaged</i>	104	133	46.6	52.6
Central Papuan				
<i>Motu</i>	106	107	59.8	60.7
Bali-Vitu				
<i>Vitu</i>	106	114	50.9	54.4
New Ireland	104.0	110.5	53.5	58.0
NW Solomonic	100.1	110.8	36.6	39.0
N Bougainville	105.0	110.4	37.9	40.6
<i>Banoni</i>	106	132	35.6	38.6
Mono-Torau	101.7	106.3	43.9	47.0
Choiseul	105.6	110.6	26.8	30.0
New Georgia	105.3	107.8	44.1	45.7
Santa Isabel	104.0	112.8	34.6	35.8
SE Solomonic	106.0	115.8	62.9	65.8
N Vanuatu/Banks-Torres				
<i>Mota</i>	106	122	63.9	66.4
Central Pacific/Fijian				
<i>Bauan</i>	106	108	67.6	71.3

the database, but this must often reflect the fact that the words collected in different languages actually have different meanings, distorting retention percentages. In the light of these weaknesses, the list of 199 meanings was reduced to the list of 106 meanings shown in the Appendix and the analysis was repeated

Table 1 summarises the analysis of retentions based on the 106-meaning list. Language names are in italics. One list, Piva, was excluded because it covered only 73 of the 106 meanings. Roman labels ('North New Guinea' etc) refer to groups of languages named in Lynch et al. (2002:ch 5). There are four numerical columns. The first shows the (average) number of meanings (out of a possible total of 106) represented in the database for each language or group. The second shows the number of entries, which is usually greater than the number of meanings because of the inclusion of alternative items. The third shows the percentage of entries which are retained from pOc: that is, they reflect a pOc etymon with (more or less) the same meaning as that etymon, as explained above. The fourth column shows the percentage of entries that reflect a pOc etymon, regardless of whether the pOc meaning is retained. This is ignored in the remaining discussion.

The difference in retention rates between NWS and SES is dramatically clear from Table 1. The average retention rate for NWS languages is 36.6%, for SES 62.9%. None of the sample languages outside the Solomons displays a retention rate anywhere near as low as NWS. The lowest is the papuanised language *Gedaged*, at 46.6% (*Motu*, also papuanised, lies at 59.8%). As Table 2 shows, the highest retention rates occur in SES (*Gela* at 73.2%, *Tolo* at 70.4%) and in the sample languages from Remote Oceania (*Bauan* Fijian at 67.66%).

Table 2 shows the analysis language by language for New Ireland, NWS and SES. This reveals

Table 2: Analysis of pOc retentions for 106 meanings in New Ireland and Solomons languages

	No. of meanings	No. of entries	Retention rate (%)	Reflexes of pOc items (%)
New Ireland	104.0	110.5	53.5	58.0
<i>Tigak</i>	101	110	51.8	56.4
<i>Tabar</i>	106	108	55.6	61.1
<i>Lihir</i>	106	110	58.2	62.7
<i>Kandas</i>	103	114	48.2	51.8
NW Solomonic	104.6	110.8	36.6	39.0
N Bougainville	105.0	110.4	37.9	40.6
<i>Nehan</i>	106	107	35.5	38.3
<i>Solos</i>	102	108	36.1	38.9
<i>Haku</i>	106	110	40.0	43.6
<i>Teop</i>	106	121	30.6	32.2
<i>Taiof</i>	105	106	47.2	50.0
<i>Banoni</i>	106	132	35.6	38.6
Mono-Torau	101.7	106.3	43.9	47.0
<i>Uruava</i>	95	97	44.3	47.4
<i>Torau</i>	104	111	44.1	47.7
<i>Mono</i>	106	111	43.2	45.9
Choiseul	105.6	110.6	26.8	30.0
<i>Varisi</i>	106	111	28.8	33.3
<i>Vaghua</i>	105	110	30.9	34.5
<i>Ririo</i>	106	110	26.4	30.0
<i>Babatana</i>	106	114	24.6	26.3
<i>Sisingga</i>	105	108	23.1	25.9
New Georgia	105.3	107.8	44.1	45.7
<i>Simbo</i>	103	108	50.0	50.0
<i>Roviana</i>	106	108	49.1	49.1
<i>Hoava</i>	106	107	37.4	39.3
<i>Marovo</i>	106	108	39.8	44.4
Isabel	104.0	112.8	34.6	35.8
<i>Kia</i>	106	112	32.1	32.1
<i>Kokota</i>	101	101	35.6	38.6
<i>Kilokaka</i>	105	111	36.0	36.9
<i>Maringe</i>	104	127	34.6	35.4
SE Solomonic	106.0	115.8	62.9	65.8
Gelic-Guadalcanal	106.0	119.3	65.3	67.2
<i>Bugotu</i>	106	138	52.2	53.6
<i>Gela</i>	106	112	73.2	75.0
<i>Tolo</i>	106	108	70.4	73.1
Malaita-Makira	106.0	114.0	61.8	65.0
<i>Lau</i>	106	132	59.1	61.4
<i>Kwaio</i>	106	108	53.7	57.4
<i>Kwara'ae</i>	106	114	63.2	66.7
<i>Toabaita</i>	106	108	61.1	64.8
<i>Sa'a</i>	106	116	65.5	69.0
<i>Santa Ana</i>	106	106	67.9	70.8

Table 3: Variation in NW Solomonic retention rates

	No. of languages	Mean retention rate (%)	Range of retention rates (%)	Extent of range
New Ireland	4	53.5	48.2 – 58.2	10.0
NW Solomonic	22	36.6	23.1 – 50.0	26.9
New Georgia	4	44.1	37.4 – 50.0	12.6
Mono-Torau	3	43.9	43.2 – 44.3	1.1
N Bougainville	5	37.9	30.6 – 47.2	16.6
<i>Banoni</i>	1	35.6	—	—
Santa Isabel	4	34.6	32.1 – 36.0	3.9
Choiseul	5	26.8	23.1 – 30.9	7.8
SE Solomonic	9	62.9	52.2 – 73.2	21.0

another difference between NWS and SES. Within the latter the highest retention rate is Gela at 73.2%, the lowest Bugotu at 52.2% (range = 21). Within NWS the highest is Simbo at 50%, the lowest Sisingga with 23.1% (range = 26.9). At first sight, it seems that the two groups have a similar profile, but that SES languages have retention rates around 23–29% higher than NWS. But a closer look reveals that this is not the whole story. Bugotu has a rather low retention rate by SES standards, the more so as its closest relative appears to be Gela, with a high 73.2%. The reason for the low rate in Bugotu is almost certainly that at Tataba village on the southeast tip of Isabel, Bugotu is spoken alongside Maringe. Tataba people are bilingual, and Bugotu has borrowed from Maringe, lowering its retention rate.¹⁶ A few such borrowings can be identified in the 106-meaning list:

- (5) a. Bugotu *kei-* ‘tooth’; cf Kok *kei-*, Kil *khe?i-*, Mge *khe?i* but Gela *livo-* < pOc *lipon
 b. Bugotu *ðehe* ‘die’; cf Kok Kil Mge *lehe* but Gela *mate* < pOc *mate
 c. Bugotu *sesehu* ‘grass’; cf Sis *sisíu*, Mvo *tsetseu*, Kia *sesehu* but Gela *yaoyabona*
 d. Bugotu *kola-* ‘liver’; cf Bab Hoa Mvo Kia Kok *kola-*, but also, Bugotu *ate-*, Gela *ate-* < pOc *qate

Table 3 summarises from Table 2 the mean retention rates and ranges of retention rates for NWS languages (New Ireland and SES are shown for comparison). NWS groups are arranged in rank order of mean retention rates. Certain points emerge:

- (6) a. The rank order of mean retention rates bears no special relationship to geographical locations. Santa Isabel and Choiseul, with low retention rates, are in the east, but so is New Georgia, with a higher rate.
 b. The highest retention rate in New Georgia is 50%, in North Bougainville 47.2%, suggesting that pNWS had a retention rate above 50%, i.e. a little higher than Kandas in Southern New Ireland with 48.2% (Kandas is the closest relative of pNWS included in the database).

The greater variation among retention rates of NWS subgroups suggests that the subgroups of NWS have more varied local histories than the two major subgroups of SES. This is supported by the

¹⁶ There are also Santa Isabel borrowings in Gela and Tolo, but apparently fewer than in Bugotu.

Table 4: Innovative lexical items in NW Solomonian groups in the 199-meaning list

	No. of languages	Mean retention rate from pOc (%)	No. of exclusively shared innovations	No. of innovations
New Georgia	4	44.1	38	77
Mono-Torau	3	43.9	11	52
N Bougainville	5	37.9	28	64
Banoni-Piva	2	35.6	12	37
Santa Isabel	4	34.6	59	92
Choiseul	5	26.8	41	91
W Choiseul	2–3	29.9	31	...
E Choiseul	2–3	23.9	24	...

observation in 6a, which implies that the differences between NWS subgroups reflect their histories more or less *in situ*. The observation in 6b that pNWS had a retention rate above 50%, reflected in the protolanguages of the New Georgia and North Bougainville subgroups, means that losses of pOc reflexes leading to lower retention rates within these subgroups must have occurred independently within each subgroup. The Choiseul and Santa Isabel subgroups, however, display much less internal lexical variation, and suffered a reduction in their retention rates early in their separate histories. Proto Choiseul seems to have had a retention rate of, say, 33%, Proto Santa Isabel of around 38%.

The ranges of retention rates in North Bougainville and New Georgia suggest a chequered history even within these groups. Taiof displays the highest retention rate within North Bougainville, explained by Lincoln’s (1976a:422) observation that Taiof has been isolated from other North Bougainville languages—and has evidently had less contact with other languages than they have.

7 Shared lexical innovations in Northwest Solomonian groups

7.1 Distribution

I have attempted to reconstruct innovative lexical items for the 199-meaning list in the various Northwest Solomonian groups in order to gain an impression of how lexical innovations are distributed among them.

The second and third columns of Table 4 are repeated from Tables 2 and 3. The fourth column shows the number of exclusively shared innovative lexical items in each group, supporting the claim that each is indeed a distinct subgroup. The fifth column shows the number of shared innovative lexical items in each group, including items reflected in more than one group, giving some sense of the degree of lexical innovation in NWS languages. ‘Innovative’ here means ‘not identifiable as Oceanic’, except in the case of the *-r-initial pronouns mentioned in §3. There is some expected inverse correlation of lexical innovations with retention rates from pOc. The Banoni-Piva figures are depressed in relation to figures for other groups because the Piva list covers only 87 meanings. Santa Isabel and Choiseul display 59 and 41 exclusively shared lexical innovations respectively (92 and 91 if we include innovations shared across group boundaries), which correlate with their low retention rates of 34.6% and 26.8%.

There is a complication in the Choiseul figures, as it became clear during the analysis that Choiseul falls lexically into two areas, West (Vagua and Varisi) and East (Babatana, Sisingga), which overlap

Table 5: Innovative lexical items exclusively shared by two or three NW Solomonic groups (199-meaning list)

	—	MT	Ch	NGe	Is
NBv and BP	7	2	1	0	1
NBv and MT	9	—	2	1	0
NBv and Ch	3	—	—	0	1
NBv and NGe	3	—	—	—	0
NBv and Is	1	—	—	—	—
BP and MT	5	—	1	0	0
BP and Ch	0	—	—	1	0
BP and NGe	3	—	—	—	0
BP and Is	0	—	—	—	—
MT and Ch	6	—	—	1	0
MT and NGe	3	—	—	—	1
MT and Is	5	—	—	—	—
Ch and NGe	13	—	—	—	5
Ch and Is	10	—	—	—	—
NGe and Is	4	—	—	—	—

Table 6: Innovative lexical items exclusively shared by four or more NW Solomonic groups (199-meaning list)

NBv/BP/TM/Ch	2	<i>*-r-o[e]</i> ‘you sg’, <i>*pisa</i> ‘three’
NBv/BP/NGe/Is	1	<i>*kolomo</i> ‘water’
NBv/Ch/NGe/Is	1	<i>*kap^wa[ta]</i> ‘skin’
MT/Ch/NGe/Is	2	<i>*siko</i> ‘steal’, <i>*kavere</i> ‘spider’
NBv/BP/TM/Ch/Is	1	<i>*siqop^wa</i> ‘intestines’
NWS	1	<i>*-r-[i]au</i> ‘I’

in Ririo, implying that there was once a dialect chain stretching the length of the island.¹⁷ If we infer, as I do below, that lexical variation is largely due to lexical borrowings from different Papuan sources, then we must conclude that in earlier times there were two rather different Papuan languages on Choiseul. In addition to the 91 innovations which include West and East Choiseul, then, Table 4 shows 31 West Choiseul and 24 East Choiseul innovations.

The differences between the figures in columns 4 and 5 of Table 4 indicate that a fair quantity of innovative items is shared between groups. This suggests the possibility that larger historic groupings of NW Solomonic languages might be identifiable on the basis of exclusively shared lexical innovations, and the relevant figures are presented in Tables 5 and 6. The first column of figures in Table 5 shows the number of exclusively shared lexical innovations in each pair of languages, and the re-

¹⁷ There is a similar division in NGe, but my database contained only one WNGe list (Simbo), so I have not analysed the West/East difference here.

maining columns show the number in each trio. If there were larger historic groupings, then we would expect Table 5 to display rather larger numbers of exclusively shared lexical innovations than it does. In comparison with the numbers of innovations defining the six groups in Table 4, the figures in Table 5 are small, and suggest that each group shares a few innovations with its neighbours, as a result either of borrowing or of the differentiation of the NWS groups out of an earlier dialect network. That is, NWS is a linkage.

A possible exception to this generalisation are the Choiseul, New Georgia and Santa Isabel groups. Choiseul and New Georgia share 13 innovations, New Georgia and Santa Isabel 10, and the three groups together 5. These innovations tell us that there is some kind of relationship between these three groups (with Mono-Torau sitting on the periphery) and that it is somewhat stronger than any other relationships apart from those within the six groups. But they do not tell us what kind of relationship it is. The shared innovations *may* indicate that the three groups have an exclusively shared common ancestor, but they may also reflect borrowing or an earlier relationship within an early NWS dialect network.

Of the sparse innovations noted in Table 6, only one, *-r-[i]au ‘I’, is reflected in all six groups and is thus unambiguously reconstructable to pNWS. It seems very likely that the others, however, are also of pNWS antiquity but have been lost in one or two of the six groups in the course of their diverse histories. The same may also be true of some of the items for which reflexes have been found in only three groups.

These findings have consequences for reconstruction. To all intents and purposes NWS consists of six coordinate subgroups which are probably the result of a quite rapid spread of pNWS speakers through the smaller offshore islands and along the coastal strips of Bougainville, Choiseul, New Georgia and Santa Isabel. The resulting network of dialects diversified into today’s subgroups as NWS speakers interacted with speakers of a variety of Papuan languages. This process entailed the replacement of Oceanic etyma by Papuan loans at varying rates, as noted in §6. In these circumstances it would be otiose to insist that an etymon be reflected right across NWS in order to reconstruct in pNWS. Instead, I assume that any item that is reflected on Bougainville (in NBv, BP or MT) and in the Western Solomons (in Ch, NGe or Is) is of pNWS antiquity. Because one may reasonably infer that pNWS was spoken on Buka, I also assume that any item reflected in NBv and at least one of BP and MT may also be reconstructed to pNWS. These criteria are loose for convenience. The looseness does not affect the arguments of this paper, but it does mean that ‘pNWS’ means ‘early NWS’ rather than exactly denoting the protolanguage.

7.2 Origins

Finding the origins of innovatory lexical items in NWS languages is no simple task. Given the two-wave proposal (§5), one might expect to find two kinds of origin:¹⁸ borrowings from Papuan languages and borrowings from Old Oceanic languages. In addition, it is very likely that some unsourced items have an as yet unrecognised MM origin.

7.3 Lookalikes

Associated with the question of origins is the phenomenon of ‘lookalikes’. Lookalikes are lexical items in different NWS languages which are similar in meaning and form but which do not display

¹⁸ There are also a few words reflecting etyma in MM interstage languages of a higher order than pNWS, for example, pMM *qase(n), *qasen-i- ‘count’ (Vitu *yaðeni-*, Tabar *ase*, Neh *ah*, Teo *ahe*, Mon *kala*, Mvo *ase*, Kil *a?ahe*); Proto New Ireland/NWS *paka ‘leaves, feather’ (Tigak *pak* ‘leaf’, Tabar *pakapaka* ‘leaf’, Teo *paka* ‘feather, leaf’, Hak, Teo *pakapaka* ‘wing’, Tai *pakpak* ‘wing’). Such items are not included in the figures in Tables 4 and 5.

regular sound correspondences. This means that they cannot be the direct result of shared inheritance, but are presumably the outcome of borrowing at some point.

Pairs of lookalikes fall into two categories. The first category consists of a lexical item directly inherited from pOc (or from an interstage later than pOc but earlier than pNWS) paired with an item that appears to be descended from the same pOc etymon but via borrowing. For example, the items in 7a are fairly regular reflexes of pOc *qalipan ‘centipede’. The minor irregularities they display are not likely to be due to borrowing.¹⁹ The items in 7b, however, reflect a putative pNWS *kalivaga, where pOc *q- is irregularly reflected as *k- and pOc *-n as *-g-. One may reasonably infer that this was borrowed into an early NWS dialect from another Oceanic language, probably an Old Oceanic language, in which pOc *q- was reflected as *k-.

- (7) a. pOc *qalipan ‘centipede’ > pNWS *alivaŋa > **NBv** Neh *hilaŋ* (metathesis), Hak *lihaŋa*, Tai *aifaŋ*, **MT** Uru *rivana* **WNGe** Lun *li-livaŋa*, Ndu *liwaŋa*, **Is** Bla *n-alifa*, Gho, Mge *n-alhiŋa*
 b. pNWS *kalivaga > **NBv** Tai *zanevaga*, **ECh** Rir Bab Sis *kaligava* (metathesis)

The probability of a chance similarity between 7b and 7a is very low indeed, given that the pOc etymon had four consonants and three vowels, each needing to be matched by chance. The question with lookalikes, of course, is the point at which chance becomes a major factor. One could establish criteria based on the concepts in Nichols (1996:50–54), but this would entail discussion beyond the scope of this paper. Intuitively, however, it is obvious that 8a and 8b, with fewer resemblant syllables, are more likely to be outcomes of chance resemblances—but may also indirectly reflect *qalipan.

- (8) a. **NGe** Mvo *lipata* ‘centipede’
 b. pNWS *kali ‘centipede’ > **NBv** Teo *kare*, **MT** Mon *ale-le*, **WCh** Var *kali-kali*

Lookalikes like the pair in 7 are based on a set of inferences similar to those made by Biggs (1965) about the history of Rotuman, where one inherited and two borrowed layers of vocabulary are distinguished, or by Ross (1996) about Yapese, where five contributing sources, one inherited and four borrowed, are identified. The difference, however, is that the layers of vocabulary in NWS are of greater antiquity, and the source language(s) of one of the layers cannot be reconstructed.

There is strong evidence to infer that NWS languages have borrowed from Papuan neighbours, and it is possible that an item like 7b was borrowed from a Papuan language which had borrowed it from an Old Oceanic source.

The second category of lookalikes consists of pairs neither of whose members can be sourced to pOc. This may mean one of several things (apart from chance resemblance). The items may be directly and indirectly descended from an as yet unreconstructed pOc etymon, or they may reflect different borrowings of a single Papuan item, or they may reflect borrowings of cognate items from two different Papuan languages. For example, pNBv *ma(l,r)oto in 9a and pNWS *manoqa in 9b, both ‘ten’, look as if there is some commonality in their history: the first syllable of both is *ma-, the second is an apical plus *-o-. But the apicals don’t correspond and the final syllables don’t match, suggesting that the commonality is mediated by borrowing, perhaps of items for ‘ten’ from different but related Papuan languages. The fact that the pOc term for ‘ten’ in 9c also survived—and its reflexes coexists with those of *ma(l,r)oto in the NBv group—reinforces the possibility that borrowings were often very localised.

¹⁹ pNWS *-ŋ- is an idiosyncratic reflex of pOc *-n (Ross 1988:223); Nehan, Haku, Uruava, Lungga and Nduke have lost initial *qa- (cf pOc *qapaRa ‘shoulder’ > pNWS *para > Hak *hala-hala*, Tor *ara*, Mon *hala*, Lun Ndu Rov *vara*, Lag *fara*); and Taiof loss of *-l- is unusual.

- (9) a. pNBv *ma(l,r)oto ‘ten’ > Sol *manot*, Pet *malot*, Hak *maloto*, Sel *malto*
 b. pNWS *manoya ‘ten’ > **BP** Ban *manoya*, **Ch** Var *manoya*, Rir *manua*, Bab *mano*, Sis *mano*, **NGe** Sim *manoya*, Lun Ndu *manoya-puta*, Rov Hoa *maneye-puta*
 c. pOc *saŋapulu ‘ten’ > PNS *saŋavulu > **NBv** Neh *haŋaulu*, Tai *safunu*, Teo *savun*, **MT** Tor *saunu*, Mon *lahulu*, Uru *avuru*, **WCh** Vag *ŋəvəl*

7.4 Borrowings from Papuan languages

I have sourced no Papuan loanwords to my satisfaction.

Three groups of present-day Papuan languages are relevant to the search for sources. Much of Bougainville is occupied by the North and South Bougainville families.²⁰ Lincoln (1976b) investigated the claim that Piva basic vocabulary includes loans from neighbouring Rotokas (N Bougainville). He found no evidence for them. The South Bougainville family falls into two subfamilies, Nasioi and Buin, represented respectively by dictionaries of Nasioi (Hurd & Hurd 1974) and Buin (alias Telei; Laycock 2003). The most likely candidate for borrowings from the Nasioi subfamily is Banoni, and from the Buin subfamily, Mono, but I have found no evidence of such loans.

The third Papuan group is the Central Solomons family, consisting of four geographically scattered languages: Bilua, Baniata, Lavukaleve and Savosavo. The first two are located in the New Georgia group and are plausible sources for borrowings in New Georgia languages. Bilua is also a plausible borrowing source for Choiseul languages. The wordlists in Tryon & Hackman (1983) show a number of borrowings involving each language, but in a number of cases the direction of borrowing is clearly from Oceanic to Papuan, and there is no unambiguous evidence of borrowing in the opposite direction.

The fact that there are so many unsourced NWS etyma is less puzzling than it may seem. A careful look at Todd’s (1975) word lists for the Central Solomons family reveals that almost the only recognisable cognates among them are Oceanic loans. The Papuan languages of northwest island Melanesia have been *in situ* for such a long time that their basic vocabularies have diverged beyond recognition. This is true not only of the Central Solomons languages, but also of the North and South Bougainville families. The three families are apparently unrelated to each other, and relationships within each family seem rather distant, a reflection of their great time depth (Ross 2001, Dunn et al. 2002). It can be readily inferred that before the arrival of Oceanic speakers in the region, there were far more Papuan languages in northwest island Melanesia than there are today, and that they already reflected a great degree of diversity. Dunn et al. (2005) have suggested that their diversity dates back more than 10,000 years. If this is so, then it is eminently likely that the present-day Papuan languages referred to in this section bore no recognisable relationship to the now lost languages that contributed vocabulary to early interstages of NWS and that many NWS lexical items will thus remain unsourced.

7.5 Borrowings from Old Oceanic languages

As noted in §7.3, there are a number of lookalikes where one member of the pair is directly inherited from pOc and the other is an indirectly inherited reflex of the same etymon, apparently borrowed from another (Old) Oceanic language. Since these lookalikes are evidence for the hypothesis that there was Old Oceanic settlement in the NWS region before the MM settlement reflected by NWS itself, the most clearly attested of them are set out below. One, 7, was given in §7.3.

Whereas the forms in 10a reflect normal pNWS lenition of pOc *-k- as *-ŋ- and loss of pOc *-w-, the forms in 10b reflect an Oceanic form in which pOc *-k- is unlenited and *-iw- is preserved as *-u-.

²⁰ Sometimes called the West and East Bougainville families.

- (10) a. pOc *bakiwa ‘shark’ > pNWS *bayea > Bab Sis *bəza*,²¹ Gha Lun Sim *bayea*, Kil Kok *bae-su*, Mge *baʔe-su*
 b. pNWS *bakuai ‘shark’ > **NBv** Neh *bakue*, Sol *bake*, Hak *baki*, Sel *buei*, Teo *baku-baku*, **BP** Ban *bakuo*, **MT** Tor *vavoi*, Uru *baku-baku*, Mon *baʔoi*, **WCh** Vag *bakui*, Var *bakuai*, Rir *boʔoei*

The forms in 11a are regular reflexes of pOc *pituqun, whereas those in 11b reflect an Oceanic form with reduplication, unlenited *p-, loss of *-q- and *-n. In 11c pECh *-putu seems to be a separate Old Oceanic reflex. In 11d pNGe *pi(no)-pino raises the classic lookalike problem: is it an Old Oceanic reflex (with loss of *-t- and retention of *-n-?) or does it resemble pNWS *pi(to)-pito by chance?

- (11) a. pOc *pituqun ‘star’ > pNWS *vitu(ɣu)nu > **MT** Uru *vesunu*, **Is** Gho Mge *na-thunu*²²
 b. pNWS *pi(to)-pito ‘star’ > **NBv** Neh *pito-pit*, Sol *bi-pit*, Hak *pito-pito*, Ban Piv *pi-pito*, Mon *vito-vito*, Tor *vi-vito*
 c. pECh *sisiri-putu ‘star’ > Bab Sis *sisiri-putu*
 d. pNGe *pi(no)-pino ‘star’ > Ndu Ugh *pi-pino*, Rov Kus *pinopino*

The forms in 12a are regular reflexes of pOc *boRok, whereas those in 12b are open to two analyses: they reflect either (i) loss of pOc *-R- and unlenited *-k-, or (ii) pOc *-R- as -k- and loss of pOc *-k-. Papuan languages have also borrowed this term from Oceanic: Bilua *bolo* appears to be a borrowing from WNGe *boroyo*, Baniata *bo* from ENGe *boko*.

- (12) a. pOc *boRok ‘pig’ > pNWS *boroyo > **BP** Ban *boroyo*, **WNGe** Lun Ndu *boroyo*
 b. pNWS *boko > **NBv** Tai *vo*, **MT** Tor *bo*, Mon *boʔo*, **Ch** Vag Var *boko*, Rir *boʔ*, Sis Bab *boko*, **ENGe** Rov Hoa *boko*

The Santa Isabel reflexes in 13a *appear* to be directly inherited, as do the Piva (BP) and Sim (WNGe) reflexes of the same form with prefixed *[ma]ma- in 13b. However, pCh *madaka- in 13c appears to be an indirectly inherited reflex of this prefixed form, with the same diachronic phonological ambiguity as pNWS *boko ‘pig’ in 12b. If pCh *madaka- is indeed indirectly inherited (i.e. an Old Oceanic loan), then the possibility exists that *[ma]ma-dara(ɣa)- in 13b and even *dara(ɣa)- in 13a are also indirectly inherited.

- (13) a. pOc *draRaq ‘blood’ > pNWS (?) *dara(ɣa)- > **pIs** Mge Kok Kil *dadara*
 b. pNWS (?) *[ma]ma-dara(ɣa)- > **BP** Piv *ma-rana-*, **WNGe** Sim *mama-dara-*
 c. pCh *madaka- > Vag Bab Sis *madaka-* Rir *madak*

I have included this example not because I am especially optimistic about the suggested analysis above but because it gives some flavour of the analytical problems that NWS lexicon confronts us with. Also found are Rov *ehara-*, Mvo Van *juka-*, Hoa Kus *mazuka-* and Kia *busaka-*, all ‘blood’. The (Papuan) Bilua term is *dara-*, surely borrowed from an Oceanic language, Old or Western, but there is no evidence of the form in the usual Oceanic sources of Bilua borrowings.

Returning to more straightforward examples, 14a contains regular reflexes, whilst 14b and 14c appear to have been borrowed from Old Oceanic languages at different points in NWS history. The former has voiced *b- and retains unlenited *-k-, and the latter has voiceless *p- and loses *-k-.

- (14) a. pOc *p^wekas²³ ‘defaecate’ > pNWS *beyasa > **NBv** Neh *beh*,²⁴ **BP** Ban *beyasa*, **WCh**

²¹ Here -z- appears to reflect *-y-: *bayea > *baea > *baya.

²² Aspirated -th reflects pls *-ft-.

²³ The reconstruction of *p^w- is uncertain: see Ross et al. (1998:16).

²⁴ Neh -h reflects *-s.

Vag *biya*, Var *beya*

b. pNBv *beka > Teo *bebeka*

c. PWS *pea > **MT** Mon *pea*, **ECh** Rir Bab Sis *pia*, **NGe** Gha Lun Ndu Rov Van *pea*

In 15 there appear to be two pNWS forms. The first reflects pOc *wakaR regularly with loss of *w-. The second reflects *w- as *b- and was apparently an Old Oceanic form.

- (15) a. pOc *wakaR ‘root’ > pNWS *ayara- > Teo *ana*, Sel *ara*, Tor *agara-*, Uru *agara-*, **NGe** Ndu *ayara*, Hoa *ayoro*, **Is** Kia Kok Kil *zagra-*
b. pNWS *bayara- **NBv** Tai *vora-*, **BP** Ban *bayara-*, Piv *bagara-*, **WNGe** Sim Lun *bayere*, **Is** Lag *bakla*

The examples above stand out because the pOc forms have three consonants and 2–3 vowels and doublet reflexes, as well as reflexes from across the NWS area, i.e. there is adequate evidence that their doublets reflect borrowings from Old Oceanic forms. There are a number of other forms that do not satisfy these criteria, and varying degrees of doubt must subsist as to whether they are reflexes of Old Oceanic forms, chance resemblances, or even irregular reflexes of MM forms. Some of these forms have indeed been analysed in the past as irregular reflexes of MM forms.

In Ross (1988:224) I remarked that pOc/pNWS *s became Proto West Choiseul²⁵ *j in some lexical items. My examples were 16 and 17. Notably pOc *siku is apparently not reflected elsewhere in NWS languages, increasing the likelihood that pWCh *jiku was a local borrowing. Both items reflect unlenited pOc *k rather than the usual inherited *γ. I return to reflexes of pOc *kusupe below.

(16) pOc *siku ‘elbow’ > pWCh *jiku²⁶ > Vag *zə-zəkə*, Varisi *zi-ziku*

(17) pOc *kusupe ‘rat’ > pWCh *kuju > Vag *kəj*, Var *kuzu*, Rir *kuj*

There are numerous other lookalikes in Choiseul languages, correlated with the fact that they have the lowest retention rates in NWS, but it is often hard to be sure that they are not chance resemblances. For example, Tryon & Hackman (1983:61) note pCh *pade ‘house’ as an irregular reflex of pOc *pale. It is likely that they are right, and that this is an Old Oceanic loan, but the possibility that it is a chance resemblance cannot be ruled out.

There are a number of apparent Old Oceanic borrowings limited to the Santa Isabel group. The doublet form in 18b, is reduplicated and displays loss of pOc *-t-.

- (18) a. pOc *mataqut ‘fear’ > PWS *matayutu > **NBv** Pet *matout*, Hak *matutu*, **NGe** Sim *matutu*, Gha Lun Rov *matayutu*
b. pIs *mamayu > Kia *mamayu*, Kil Mge *hmayu*

The items in 19 and 20 are given by Tryon & Hackman (1983:61) as illustrations of the Santa Isabel reflex of pOc *w, but there is good evidence that pOc *w was lost in the Santa Isabel group, as in other NWS languages (Ross 1986:186–197). It thus seems likely that Santa Isabel items like these, in which pOc *w has merged with *p, represent Old Oceanic loans (this does not belie Tryon and Hackman’s basic point, that these items are uniquely shared innovations of Santa Isabel languages).

(19) a. pOc *ma-ñawa ‘breathe’ > pNWS *ma-ñaa ‘heart’ > **Is** Gho *ma-ñaa* ‘heart’

²⁵ Labelled ‘Proto Choiseul’ by Ross 1988.

²⁶ Tryon & Hackman (1983:60) points out that when an initial consonant is lost before *-i- in Choiseul, z-accretion occurs in all languages except Vaghua. Here, however, z- also occurs in Vaghua, indicating that this is not accretion.

- b. PWS *ma-ñava ‘heart’ or ‘liver’ > **WCh** Vag *ma-nava* ‘liver’, **Is** Kok Bla *na-nafa* ‘heart’, Lag *na-nafa* ‘heart, breathe’, Bla *fahaye-nafa* ‘breathe’, Gho *ña-ñafa* ‘breast’, Kok *na-nafa* ‘breathe’, Mge *na-ñafa* ‘heart’
- (20) a. pOc *siwa ‘nine’ > pNWS *sia > **NBv** Neh *lu-sio*, Sol *sie*, Hak *to-si*, Tai *sia*, **MT** Tor *sia*, Mon *u-lia*, Uru *ia*, **Ch** Var *ka-ia*, Rir *zia*, Bab *zia*, Sis *zia*, **NGe** Sim Rov Hoa *sia*,
- b. pIs *n-heva > Ko *ηheva*, Kil *hneʔva*, Mge *hnevai*

The two lookalike sets below further exemplify the difficulty of unravelling NWS lexical history. pNWS *tuyur, the regular reflex of pOc *tuqur, is well attested in 21a. It is tempting to assume that pMT *tegese in 21b, pCh *deyere in 21c and pIs *tetu in 21d are all borrowings of Old Oceanic reflexes of pOc *tuqur, but again they may be chance resemblances.

- (21) a. pOc *tuqur ‘stand’ > pNWS *tuyur > Neh *tur*, Sol *tonon*, Teo *sun*, Tai *tutun*, Ban *tsuyonu*, Uru *toru* Gha Lun Sim Ndu Rov *turu*
- b. pMT *tegese > Mon Tor *tegese*
- c. pCh *deyere > Var Vag *deyere*, Rir *der*, Sis *dire*, Bab *dere*
- d. pIs *tetu > Kok Kil Kia *tetu*

The forms for ‘rat’ provide even greater difficulties. We know that pOc *kusupe ‘rat’ was reflected as pMM *kusuve, i.e. with an unlenited initial *k-, as reflected in the MM forms in 22 from languages outside NWS.²⁷

- (22) pOc *kusupe ‘rat’ > pMM *kusuve > **Bali-Vitu** Vitu *kuvuðe* (*s/*p metathesis), Bali *kuvuzeke* (*s/*p metathesis), **Willaumez** Bulu Bola *kuruve*, Nakanai *kusuke*, **Tabar** Lihir *kues*, **Madak** Lamasong Madak *kisap*, Barok *kisuo*, **S New Ireland** Konomala *kusi*, Minigir *kusuva*, Label Siar *kusup*, Bilur *kue*, Kandas *kusupu*, Ramoaina *kaupa*

The expected pNWS form would also have been *kusuve, but there are no forms with a reflex of *-v-, and so I reconstruct pNWS *kusu(v)e. Proto North Bougainville *kiso in 23a is one of several lexical items in which pOc *u became pNBv *i (Ross 1988:223), and is quite possibly directly inherited. Proto Santa Isabel *kusi in 23c is also a plausible directly inherited reflex (cf Konomala *kusi* in 22), which has also been borrowed as Bugotu, Gela (SES) *kuhi*.

- (23) pOc *kusupe ‘rat’ > pMM *kusuve > pNWS *kusu(v)e >
- a. pNBv *kiso > Neh *kih*, Sol *kiso*, Hak *isu*, Teo *kuho*, Tai *kiso*,
- b. pBP *kiso > Ban *kiso*
- c. pIs *kusi > Mge *na-khusi*

Lookalikes are listed in 24. Proto Mono-Torau *kuake was perhaps an Old Oceanic borrowing, in which pOc *-s- has been lost and *-p- had been replaced by *-k- (cf Nakanai *kusuke* in 22). Proto New Georgia *karuje is not readily explicable, but pECh *ruji is plausibly a version of *karuje with first-syllable loss. Proto East Choiseul *kuju was discussed above. It is possible that pENGe *kutu was a borrowing of an Old Oceanic form cognate with the source of pECh *kuju. However, similar forms are found in New Ireland (Tabar *kotu*, Tangga *kut*) and it is also possible that pENGe *kutu was cognate with them. This reminds us, incidentally, that New Ireland may also have had an Old Oceanic period, also reflected in lookalikes.²⁸

²⁷ The pSES form, attested without change in West Guadalcanal and Talise was *γusuve with lenited *γ-.

²⁸ Cf. Tigak *kusia*, Tiang *kuse*, Kara *kuf*, all ‘rat’.

- (24) a. pMT *kuake > Mon Tor *kuake*, Uru *kue*
 b. NGe *karuje > Sim *karuje*, Ndu Rov *kurezu*
 c. pECh *ruji > Bab *ruji*, Sis *roji*
 d. pWCh *kuju > Vag *kəj*, Var *kuzu*, Rir *kuj* (repeated from 17)
 e. pENGe *kutu > Hoa Mvo *kutu*

The variation in terms for ‘rat’ is so great that I have also wondered whether they represent later borrowings after MM/NWS settlement. However, this would presuppose that *Rattus exulans* followed Oceanic speakers into the region. As Matthew Spriggs comments (pers. comm., 2007), ‘the route for *Rattus exulans* into the Pacific was clearly as stowaways or snack lunches on Lapita canoes ... The other complication is of course the rich endemic rat fauna of the main Solomons compared to the Bismarck Archipelago ...’ A more likely explanation of the terms above is that they represent parallel borrowings from Old Oceanic languages.

What conclusions can we draw from the sample of lookalikes in this section? Examples 7 and 10–15 are generally more convincing evidence for the erstwhile presence of Old Oceanic languages in the NWS region than 7.5–24. There is a methodological difficulty here. The probability of chance resemblances is less with regard to three- and four-syllable forms than it is with regard to two-syllable forms—but the majority of pOc roots, at least in more basic vocabulary, had only two syllables. Reinforcing the case for Old Oceanic requires a large number of examples, and these will only be forthcoming, if ever, when detailed dictionaries are available for a good sample of NWS languages. For the moment we can say simply that the evidence favours an Old Oceanic presence in the region, and that certain pNWS etyma appear to be Oceanic but not to be directly inherited from pOc. I have also suggested that there were later, local borrowings from Old Oceanic sources, particularly in Choiseul and Santa Isabel, but this suggestion in particular requires more research.

8 Conclusions

I have shown that retention rates in NWS languages are lower than elsewhere in Western Oceanic and considerably lower than in SES languages. It is clear from the differing retention rates of NWS groups, however, that the groups must have had rather different histories. This is confirmed by the relative absence of both morphosyntactic and lexical innovations defining larger groupings within NWS, and the resultant need to recognise six primary groups within NWS.

In what respect did the histories of these six groups differ? The answer would seem to be that their speakers have all been in contact with speakers of other languages and been bilingual in those languages at various periods in their histories. Differing retention rates reflect different degrees of contact or different social relationships with speakers of other languages. Frustratingly, we have virtually no clear evidence for loans from Papuan languages, but simply large numbers of etyma for which there seems to be no other explanation. We do have some evidence, however, for loans from Old Oceanic languages.

One intriguing feature of the data and reconstructions in §7.5 is that different Oceanic reflexes occur cheek by jowl with one another (and with apparent Papuan loans) in languages of the same group. The only reasonable explanation I can offer for this is that even quite local contact histories varied. This seems clearly to have been true of west and east Choiseul and of west and east New Georgia, as well as of the three groups on Bougainville. Indeed, the thought with which I would like to finish this paper is simply that there is huge lexical diversity among NWS languages, despite their evident genealogical unity, and this must be due largely to differing contact histories. This claim will stand, I think, even if some of the detailed proposals here are replaced by better ones.

References

- Bellwood, Peter, James Fox & Darrell Tryon, eds, 1995. *The Austronesians: historical and comparative perspectives*. Canberra: Department of Anthropology, Research School of Pacific and Asian Studies, Australian National University.
- Biggs, Bruce G., 1965. Direct and indirect inheritance in Rotuman. *Lingua* 14:383–415.
- Blust, Robert A., 1981. Variation in retention rate among Austronesian languages. Paper presented to the Third International Conference on Austronesian Linguistics, Denpasar.
- , 2000. Why lexicostatistics doesn't work: the 'universal constant' hypothesis and the Austronesian languages. In Colin Renfrew, April McMahon & Larry Trask, eds, *Time depth in historical linguistics*, 311–331. McDonald Institute for Archaeological Research.
- Dunn, Michael, Ger Reesink & Angela Terrill, 2002. The East Papuan languages: a preliminary typological appraisal. *Oceanic Linguistics* 41:28–62.
- Dunn, Michael, Angela Terrill, Ger Reesink, Robert A. Foley & Stephen C. Levinson, 2005. Structural phylogenetics and the reconstruction of ancient language history. *Science* 309:2072–2075.
- Dutton, Tom, 1994. Motu-Koiarian contact, Papua New Guinea. In Dutton & Tryon 1994, 181–232.
- Dutton, Tom & Darrell Tryon, eds, 1994. *Language contact and change in the Austronesian world*. Berlin: Mouton de Gruyter.
- Geraghty, Paul, 1990. Proto-Eastern Oceanic *R and its reflexes. In Jeremy H. C. S. Davidson, ed., *Pacific Island languages*, 51–93. London and Honolulu: School of Oriental and African Studies, University of London and University of Hawaii Press.
- Green, Roger C., 1995. Linguistic, biological and cultural origins of the initial inhabitants of Remote Oceania. *New Zealand Journal of Archaeology* 17:5–27.
- , 2003. The Lapita horizon and traditions: Signature for one set of Oceanic migrations. In Christophe Sand, ed., *Pacific archaeology: assessments and prospects*, 95–120. Nouméa: Service des Musées et du Patrimoine de Nouvelle-Calédonie.
- Hurd, Conrad & Phyllis Hurd, 1974. Nasioi dictionary. Computer file (typed 1984), Summer Institute of Linguistics, Ukarumpa.
- Kirch, Patrick V., 1997. *The Lapita peoples: Ancestors of the Oceanic world*. Oxford: Blackwell.
- Kirch, Patrick V. & Terry L. Hunt, 1998. *Archaeology of the Lapita cultural complex: a critical review*. Seattle: Burke Museum. (*Thomas Burke Memorial Washington State Museum Monograph* 5).
- Laycock, Donald C., 2003. *A dictionary of Buin, a language of Bougainville*. Canberra: Pacific Linguistics. Edited by Masayuki Onishi.
- Lincoln, Peter C., 1976a. Austronesian languages: Bougainville Province. In S.A. Wurm, ed., *New Guinea Area languages and language study*, vol. 2, 419–440. Canberra: Pacific Linguistics.
- , 1976b. Banoni, Piva And Papuanization. In *Papers in New Guinea linguistics No. 19*, 77–105. Canberra: Pacific Linguistics.

- Lynch, John, Malcolm Ross & Terry Crowley, 2002. *The Oceanic languages*. Richmond: Curzon Press.
- Nichols, Johanna, 1996. The comparative method as heuristic. In Mark Durie & Malcolm Ross, eds, *The comparative method reviewed: regularity and irregularity in language change*, 39–71. New York: Oxford University Press.
- Palmer, Bill, 2002. 'Possessive' marked verbal constructions in eastern Bougainville. Paper presented at the Ninth International Conference on Austronesian Linguistics, Canberra.
- , 2003. Owners into actors: How possessive morphology became subject agreement in the languages of the Bougainville region. Unpublished paper, University of Leeds.
- Pawley, Andrew, 1975. The relationships of the Austronesian languages of Central Papua. In T.E. Dutton, ed., *Studies in languages of central and south-east Papua*, 3–106. Canberra: Pacific Linguistics.
- , 2003. Locating Proto Oceanic. In Ross et al. 2003, 17–34.
- , 2005. Where and when was Proto Oceanic spoken? linguistic and archaeological evidence. Unpublished paper, Research School of Pacific and Asian Studies, The Australian National University.
- Pawley, Andrew & Malcolm Ross, 1993. Austronesian historical linguistics and culture history. *Annual Review of Anthropology* 22:425–459.
- , 1995. The prehistory of Oceanic languages: a current view. In Bellwood et al. 1995, 39–74.
- Reid, Lawrence A., 1978. Problems in the reconstruction of Proto-Philippine construction markers. In S.A. Wurm & Lois Carrington, eds, *Second International Conference on Austronesian Linguistics: proceedings*, vol. Fascicle 1: *Western Austronesian*, 33–66. Canberra: Pacific Linguistics.
- Ross, Malcolm, 1977. Relationships of the Austronesian languages of the Sepik and western Madang coast of New Guinea. Unpublished mimeograph, University of Papua New Guinea.
- , 1982. The development of the verb phrase in the Oceanic languages of the Bougainville region. In Amran Halim, Lois Carrington & S.A. Wurm, eds, *Papers from the Third International Conference on Austronesian Linguistics, 2: Tracking the Travellers*, 1–57. Canberra: Pacific Linguistics.
- , 1986. Towards a classification of the Oceanic languages of Bougainville and the western Solomons. In Paul Geraghty, Lois Carrington & S. A. Wurm, eds, *FOCAL II: Papers from the Fourth International Conference on Austronesian Linguistics*, 175–200. Canberra: Pacific Linguistics.
- , 1988. *Proto Oceanic and the Austronesian languages of western Melanesia*. Canberra: Pacific Linguistics.
- , 1994. Areal phonological features in north central New Ireland. In Dutton & Tryon 1994, 551–572.
- , 1996. Is Yapese Oceanic? In Bernd Nothofer, ed., *Reconstruction, classification, description: festschrift in honor of Isidore Dyen*, 121–166. Hamburg: Abera.

- , 1997. Social networks and kinds of speech-community event. In Roger M. Blench & Matthew Spriggs, eds, *Archaeology and language*, vol. 1, 209–261. London: Routledge.
- , 2001. Is there an East Papuan phylum? Evidence from pronouns. In Andrew Pawley, Malcolm Ross & Darrell Tryon, eds, *The boy from Bundaberg: Studies in Melanesian linguistics in honour of Tom Dutton*, 301–321. Canberra: Pacific Linguistics.
- , 2005. Pronouns as a preliminary diagnostic for grouping Papuan languages. In Andrew Pawley, Robert Attenborough, Jack Golson & Robin Hide, eds, *Papuan pasts: cultural, linguistic and biological histories of Papuan-speaking peoples*, 15–66. Canberra: Pacific Linguistics. (*Pacific Linguistics* 572).
- , 2006. Clues to the linguistic situation in Near Oceania before agriculture. Paper presented at the symposium *Historical linguistics and hunter-gatherer populations in global perspective*, Max Planck Institute for Evolutionary Anthropology, Leipzig.
- Ross, Malcolm, Andrew Pawley & Meredith Osmond, eds, 1998. *The lexicon of Proto Oceanic*, vol. 1: *The culture and environment of ancestral Oceanic society*. Canberra: Pacific Linguistics. (*Pacific Linguistics* C-152).
- , eds, 2003. *The lexicon of Proto Oceanic: The culture and environment of ancestral Oceanic society. 2: The physical world*. Canberra: Pacific Linguistics. (*Pacific Linguistics* 545).
- Sheppard, Peter J. & Richard Walter, 2006. A revised model of Solomon Islands culture history. *Journal of the Polynesian Society* 115:47–76.
- Spriggs, Matthew, 1995. The Lapita culture and Austronesian prehistory in Oceania. In Bellwood et al. 1995, 112–133.
- , 1996. Early agriculture and what went before in Island Melanesia: continuity or intrusion. In David R. Harris, ed., *The origins and spread of agriculture and pastoralism in Eurasia*, 524–537. London: University College London Press.
- , 1997. *The Island Melanesians*. Oxford: Blackwell.
- Terrell, John E. & Geoffrey J. Irwin, 1972. History and tradition in the northern Solomons: an analytical study of the Torau migration to southern Bougainville in the 1860s. *Journal of the Polynesian Society* 81:317–347.
- Terrill, Angela, 2002. Systems of nominal classification in East Papuan languages. *Oceanic Linguistics* 41:63–88.
- Todd, Evelyn M., 1975. The Solomon language family. In S.A. Wurm, ed., *New Guinea area languages and language study*, vol. 1, 805–846. Canberra: Pacific Linguistics.
- Tryon, Darrell & B.D. Hackman, 1983. *Solomon Islands languages: an internal classification*. Canberra: Pacific Linguistics.
- Wickler, Stephen, 2001. *The prehistory of Buka: a stepping stone island in the northern Solomons*. Canberra: Department of Archaeology and Natural History and Centre for Archaeological Research, The Australian National University. (*Terra Australis* 16).

Appendix: Reduced meaning list, with percentage Proto Oceanic retentions for each meaning (ID numbers match those used in the Austronesian Basic Vocabulary Database)

1	hand	56.1	56	child	27.3	124	sea, salt	57.7
2	left	48.7	59	mother	61.4	128	sky	18.2
3	right	43.6	60	father	84.4	129	moon	35.0
4	leg/foot	40.5	61	house	52.4	130	star	25.0
5	to walk	30.6	63	name	75.6	133	rain	26.8
6	road/path	60.0	64	to say	6.3	134	thunder	55.6
7	to come	87.8	65	rope	39.6	135	lightning	9.8
11	dust	35.0	74	to kill	22.2	143	fire	19.5
12	skin	35.7	75	to die	72.5	148	white	4.4
14	belly	33.3	76	to be alive	61.5	149	red	26.1
15	bone	45.0	78	to cut	15.5	150	yellow	32.6
16	intestines	16.7	79	stick/wood	58.1	151	green	23.3
17	liver	50.0	90	to dig	39.5	153	big	23.9
18	breast	86.7	95	to fall	29.8	155	long	23.8
22	to fear	27.9	96	dog	15.9	160	painful	22.9
23	blood	50.0	97	bird	39.0	163	new	86.0
24	head	52.4	98	egg	30.0	165	bad	45.5
27	nose	72.5	99	feather	34.9	167	night	69.0
28	to breathe	38.1	101	to fly	46.3	170	when?	56.1
29	to sniff	29.5	102	rat	25.6	180	far	40.5
31	tooth	62.5	103	meat/flesh	47.6	181	where?	75.6
32	tongue	87.5	105	tail	17.9	182	I	92.5
33	to laugh	5.0	106	snake	30.0	183	you SG	67.5
34	to cry	61.9	108	louse	82.5	184	he/she	59.5
35	to vomit	84.6	109	mosquito	23.3	185	we EXC	81.6
36	to spit	68.3	110	spider	15.6	185	we INC	78.9
37	to eat	56.3	111	fish	67.5	186	you PL	85.0
40	to drink	32.6	113	branch	31.7	187	they	33.3
41	to bite	51.1	114	leaf	50.0	188	what?	79.1
43	ear	75.6	115	root	46.8	189	who?	75.6
44	to hear	78.0	117	fruit	59.5	197	one	72.1
45	eye	92.5	118	grass	12.5	198	two	87.8
48	to sleep	22.7	119	earth/soil	31.7	199	three	64.3
53	person	38.6	120	stone	48.8	200	four	75.6
54	man/male	52.4	121	sand	31.0			
55	woman/female	27.5	122	water	29.3			