

LANGUAGE & LINGUISTICS IN MELANESIA

Journal of the Linguistic Society of Papua New Guinea ISSN: 0023-1959
Vol. 44, 2026



The phonological history of 17 varieties in western Espiritu Santo (Vanuatu)

Tihomir Rangelov

Max Planck Institute for Evolutionary Anthropology, Leipzig
tihomir_rangelov@eva.mpg.de

The phonological history of 17 varieties in western Espiritu Santo (Vanuatu)

Tihomir Rangelov

Max Planck Institute for Evolutionary Anthropology, Leipzig

tihomir_rangelov@eva.mpg.de

Abstract: Espiritu Santo Island in Vanuatu is one of the world's language diversity hotspots, while also being severely underdocumented. This paper discusses the phonological history of 17 varieties of the West Santo clade, based on new data. The most dramatic change involves the reorganisation of the varieties' plosive inventories. Loss of secondary labiovelarisation and a historical linguolabial shift are attested in some varieties. The velar nasal merged with /n/ in most varieties. An idiosyncratic *s > h change likely acts as a shibboleth in two varieties. The results of this investigation advance our understanding of the history of the languages in this part of Santo and their historical relationships, although more robust subgrouping hypotheses are pending further analysis and data collection.

Keywords: Vanuatu, Santo, historical phonology

1. Introduction

The largest island of Vanuatu, Espiritu Santo (also commonly referred to as Santo), is among the most linguistically diverse places in the world while also being severely underdocumented (Crowley 2000: 50; François et al. 2015: 8; Rangelov et al. 2025; Rangelov 2026a). This article presents the phonological history of 17 varieties of the West Santo clade based on newly collected lexical data. It follows an established tradition in discussing the phonological histories of Vanuatu languages in this journal (cf. Lynch 2019a, 2019b, 2020a; Barbour et al. 2022; Barbour 2024¹). This is the first such article that discusses varieties from Santo.

The data come from 17 locations: 12 mostly geographically contiguous settlements on the west coast of Santo and five settlements on the eastern side of the high Santo Mountain Chain that dominates the western half of the island. A few of the settlements correspond to locations for which there is data and analysis in Tryon (1976), who conducted a thorough linguistic survey of Vanuatu at the time. Section 2 offers details about the relevant locations and speaker communities. The data and methods are described in §3. The paper then turns to a discussion of the reflexes of Proto North-Central-Vanuatu (Clark 2009a) sounds in the varieties at hand, starting with vowels (§4), and then proceeding to consonants in §5. Section 6 concludes the paper.

¹ Another such article is Lynch (2016), which discusses the phonological history of the Iaai language of New Caledonia. All papers cited in the above list discuss languages from Malekula, Vanuatu's second largest island, which, just as Santo, stands out for its linguistic diversity, but is better documented (Rangelov et al. 2025).

=====

This article presents results from a project aimed to survey and study western Santo's languages. Here, I focus exclusively on the varieties' phonological history. These results provide the basis for upcoming work exploring the area's linguistic diversity and the classification of western Santo's languages.

2. The varieties in the sample

The geographic locations of the 17 varieties included in this study appear on Map 1. The choice of settlements from which to collect data was primarily based on practical considerations. Given the geography of the region and the lack of infrastructure, data collection was shaped by travel arrangements, the weather, existing contacts, and resource availability. The labels have the following format [variety index]_[geographic location (village/hamlet)]_[variety endonym]. In the rest of the paper, I often refer to varieties by their indexes for brevity.

In this part of Santo, people occupy villages that rarely exceed 100-200 inhabitants, and sometimes small hamlets (*stesen* in Bislama) comprising a handful of households, usually a single extended family that recently migrated from a nearby village to its ancestral land. In the north of the relevant area (on the Cumberland Peninsula) all settlements are close to the coast. Inland settlements can be found further south, starting with the Sauriki valley on the western slopes (e.g. varieties 11,12), and, on the eastern slopes of the mountains, south of the historical location of variety 17, including where varieties 13 and 14 are spoken (this inland area is colloquially known as Big Bay Bush). Between varieties 01 and 12, data were collected from all villages and most hamlets, with the exception of the settlements Peniel, Nampeko, Ravlepa, Sulesai and Vasalea (each with fewer than 100 inhabitants). Data were not collected from these locations mostly due to logistical issues, but people from different communities agreed that in Peniel, Nampeko and Ravlepa the L1 is “exactly the same” as 01_Nokuku_Tanokuk, while Sulesai and Vasalea both speak “exactly the same” variety as 05_Tasmate_Oa. These claims remain to be verified by linguistic data. Thus, with the caveats outlined above, the data set contains data from all geographically contiguous *lanwis* (cf. next paragraph) on the west coast from Nokuku village to the Sauriki valley. On the eastern side of the mountains, varieties 13-17 are mostly not geographically contiguous. These varieties were mostly chosen in order to verify reports from west coast residents about their varieties having historical connections with these eastern varieties (see below).

The following paragraphs offer some details on the social context in which each variety is used. These data were provided during sociolinguistic interviews with the recorded speakers and other speakers (see supplementary materials for details). Most of the details about the west coast varieties are based on the data already presented by Taman et al. (2025), where more details can be found. I employ the Bislama term *lanwis* (from English *language*) to refer to speakers' perceptions about linguistic boundaries, which is related to marking identity and belonging to a place (Lindstrom 1994, Ridge 2025).



Map 1. A map of Santo and larger neighbouring islands with the varieties included in this study. In grey are relevant settlements and other locations mentioned in this article. Variety 17 is placed at its historical location (see below for details) and in the rest of the paper, variety 17 will appear at this location to account for its relevance, but also for tidier geographic presentation of the data. The index numbering starts from the northwesternmost variety and follows the west coast south to then continues in a rough south-to-north direction on the eastern side of the mountains.

01_Nokuku_Tanokuk was recorded in Nokuku village, which has 122 inhabitants. This variety reportedly belongs to one of the largest *lanwis* of the area, with around 850 L1 speakers spread over settlements to the north of Nokuku and Penouru, Nampeko and Ravlepa to the south. It was previously also labelled as the Nokuku or Nogugu language (Glottocode: noku1237, ISO 639-3 code: nkk) and there are some materials in it from the end of the 19th century and the beginning of the 20th century, as the first Christian mission in the area was established in Nokuku. Nokuku data are also included in Tryon (1976). Based on these materials, this language was also included in Clark's (2009a, 2009b) studies. It serves as a regional lingua franca of northwestern Santo. Nokuku residents report that their ancestors have occupied the coastal area for many centuries, unlike many other villages in the area, which were formed after migrations from higher altitudes.

02_Penouru_Lela, also spelled Lela' [lela?], is a dormant variety that used to be the L1 of the Penouru community (current population: 89), which moved to its current location from the nearby hills in 1959 after an epidemic killed all but 17 community members. Following the community's Christianisation, Lela' speakers shifted to Tanokuk. The last fluent speaker of Lela' died in 2022. The data were collected from rememberers of this variety.

03_Petawata_Hurahi is the L1 of the 99 inhabitants of Petawata, whose ancestors moved to the current location from settlements higher in the nearby hills. All Hurahi speakers are also fluent in the regional *lingua franca* Tanokuk.

04_Wunavae_Takani is the L1 of the 338 inhabitants of Wunavae, whose ancestors migrated to the current location from higher altitude in the 20th century.

05_Tasmate_Oa, also spelled 'Oa [ʔoa], is the L1 of most of the 148 members of the Tasmate community and reportedly also of the 168 inhabitants of Vasalea and Sulesai further north, from which I have not collected data (see above). Tasmate data also feature in Tryon (1976) corresponding to a language labelled as Tasmate and Meri (Glottocode: tasm1246, ISO 639-3 code: tmt), and is also included in Clark's (2009b) analysis.

06_Wunaruu_Joa (also known as Japwele) is the L1 of 15 people from the same family who live in Wunaru'u, where they moved in 1998, and maintain connections with Big Bay communities which reportedly speak the same *lanwis* (known under the endonym Mojoa) and live in Malau on the Big Bay coast and a settlement with an unknown name in Big Bay Bush (where the endonym Nojoa is used).

07_Elia_Jo is the L1 of 64 people living in the northern half of the village Elia. Their ancestors migrated from the hills to the current location in the late 1970s. The southern half of the village is occupied by speakers of variety 08, who migrated from Kerepua. A family of Jo speakers moved to their ancestral land in Big Bay Bush (i.e. on the eastern slopes of the mountains) to a settlement called Panas (variety 13) in the 2010s driven away by the frequent droughts on the west coast.

08_Kerepua_Tuijo and **09_Wusi_Kula** are the L1s of the 143 inhabitants of Kerepua and the 118 inhabitants of Wusi. Half of the population of Elia reportedly also speaks variety 08 as L1, although I have not collected data from that communalect. The inhabitants of Kerepua and Wusi

claim that their varieties are virtually identical, but they maintain some shibboleths, such as the linguistic endonyms *tuijo* and *kula*, both of which mean the negative answer ‘no’ in their respective varieties. The residents of both villages migrated to the coast from the hills adjacent to Mt. Tabwemasana around the middle of the 20th century. Kerepua residents report that they have close relatives in the villages Narata and Maljia on the eastern slopes of the mountains, where the same variety is spoken; I have not collected data from these villages. Tryon (1976) has three wordlists from Kerepua and Wusi and groups them as the same language (Glottocode: wusi1237, ISO 639-3 code: wsi). This language is also included in Clark’s (2009b) study. According to Tryon (2010: 288), *filwoka* Frank Lengki remembers by name four extinct varieties spoken in this area – Latu, Mwe’ea, Sinie and Menie – but nothing more is known about them.²

10_Linduri_Tie is the L1 of the 79 inhabitants of Linduri village, whose ancestors reportedly migrated from inland areas around 1975 when parts of the same community migrated to Big Bay Bush and now reside in the villages Lawuswus, Wuniu and Winsao (the latter corresponding to variety 14), where the same *lanwis* Tie is spoken. Varieties grouped under the name Tie (corresponding also to the linguonyms Navut and possibly Bura and Sinia) have been reported in earlier works and assigned the Glottocode navu1237 and ISO 639-3 code nsw. Tryon’s (1976) variety Navut is likely closely related to this variety. It is worth noting that Tryon’s data were collected before the reported migrations.

11_Sauriki_Tiale was recorded at a village called Mataivura (a.k.a Matae) in the valley of the Sauriki river. Mataivura is one of a few villages close to each other along the river banks. During data collection at Mataivura, speakers from different settlements were present and I have been assured that all settlements in the valley now speak 11_Sauriki_Tiale as L1, except for the easternmost village, Narumaji, where a variety called Akei is spoken.³ I was told that varieties very similar to 11, which also share the endonym Tiale, are spoken in the valleys of the Pareo and Kerewai rivers further south, and also on the eastern slopes of the mountains. I am not aware of data from these places and uncovering the linguistic variation in this part of Santo is a topic that warrants further empirical investigation. A language called Tiale/Malmariv is known under the Glottocode tial1239 and ISO 639-3 code mnl. Tryon’s (1976) wordlist labelled Malmariv (and recorded in Big Bay Bush) likely belongs here. This language is also included in Clark’s (2009b) study.

12_Sauriki_Kitano is a dormant variety. The data were collected from rememberers who live in the Salalia village in the Sauriki valley, and are L1 speakers of variety 11. They consider Kitano a separate *lanwis* from variety 11. Data were collected on location in Mataivura. Another variety that is reported to have been spoken in the Sauriki valley is known as Keanu but I have not found any rememberers of it.

² Sinie may or may not correspond to the Sinia linguonym, see next paragraph.

³ Data from Narumaji has not been collected yet. Akei is the name commonly used for a language spoken by a few thousand people in the southwestern corner of Santo, comprising several dialects (Tryon 2010). The Narumaji variant is likely part of this dialect chain. However, see Footnote 4 for conflicting evidence.

13_Panas_Jo is the L1 of the 15 inhabitants of the Panas hamlet in Big Bay Bush, who relocated from Elia in 2015 (see above) and keep close ties with the Elia community. They report they speak the same variety as 07. They also report some competence in 05,10,11,14.

14_Winsao_Tie is the L1 of around 225 people living in the village of Winsao. The village was settled in the 1970s by people who migrated from higher altitude areas around Mt. Santo Peak (the second highest peak on Santo, located to the south of Mt. Tabwemasana). Winsao people claim that the same *lanwis*, Tie, is spoken in other, smaller communities spanning all three sides of the mountains: on the eastern slopes of the mountains in the villages/hamlets Papaeone, Wuniu, Ora, Marama, Vora; in Linduri on the west coast (variety 10) and nearby Narumaji in the Sauriki valley⁴; in Najara in southern Santo. Winsao speakers report varying degrees of competence in 11,16,17 and Mojoa (a variety they claim is spoken in the villages Tuwut and Narata, the latter reported as speaking the same *lanwis* as 08_Kerepua_Tuijo by Kerepua people).

15_Petel_Na is the L1 of the fewer than 300 people living in Petel village and adjacent Pelwus village, which has a similar population. The ancestors of the Petel people migrated closer to the coast from adjacent higher-altitude areas around the 1980s. Petel people report that the same *lanwis* is spoken in Pialulup, Vas, Piamato, Piapurea and Pwirtike villages and adjacent hamlets, all to the north of Petel/Pelwus, with some minor variation among them. The total number of speakers of this *lanwis* may reportedly amount to around 1000. Speakers of this variety report some competence in 16_Piamatsina_Pwelmau and Pesena to the north and 17_Portolry_Tolomako and other coastal Big Bay varieties to the south. They also report comprehension of 04_Wunavae_Takani on the west coast due to historical ties. However, they report they normally switch to Bislama when communicating with people from locations further away like Hokua (by Cape Cumberland in the north), 14_Winsao_Tie or 05_Tasmate_Oa. The endonym *Na* means ‘this (one)’ in this variety.

16_Piamatsina_Pwelmau is the L1 of the over 200 residents of Piamatsina village. They report that the same *lanwis* is also spoken in nearby Pialalau (pop. 100+) and Piavu (pop. 100+). The Piamatsina community was forced by the authorities to migrate to the current location on the coast from the nearby hills in 1967. They report their variety belongs to the same *lanwis* as 03_Petawata_Hurahi, due to shared ancestry before migration took place. Piamatsina speakers report some competence, due to similarity, in the varieties spoken in Piamato and Pwirtike, and less competence in the Pesena variety to the north. When communicating with speakers of Big Bay coastal varieties to the south (e.g. Jereviu, Malau, Matantas and Port Olry, see next paragraph), they report using a variety of the Tolomako language, which they perceive as a local Big Bay *lingua franca* (*kampani lanwis* or *Bislama blong Big Bay* in Bislama). Tryon (1976) has a wordlist from Piamatsina and listed it as a separate language (Glottocode: piam1242, ISO 639-3 code: ptr).

17_Portolry_Tolomako is the L1 of roughly half of the population of Port Olry, a large village on the eastern coast of Santo, which is populated by two distinct linguistic groups: the

⁴ Note that speakers of 11_Sauriki_Tiale reported that Narumaji was an Akei-speaking village, rather than a Tie-speaking village, see Footnote 3. I did not visit Narumaji and it remains to be confirmed what its linguistic situation is.

speakers of the Sakao language (cf. Touati 2014) and speakers of the variety at hand here. The two belong to different primary subgroups of the Espiritu Santo clade (Hammarström et al. 2026). The speakers of 17_Portolry_Tolomako migrated to the present location on the east coast in the 1920s (Monnier 1994: II) when the Catholic mission of Tolomako (a village in the southwestern corner of Big Bay) moved to the Port Olry area. This Port Olry variety is thus normally classified as part of a larger language labelled Tolomako (Glottocode: tolo1255; ISO 639-3: tlm), which is also currently spoken along the coast of Big Bay from Matantas (in the southeastern corner) to the southwestern corner and around ten kilometres further up the coast to the north to at least a village called Jureviu (Clark 2006, Meyerhoff 2009). However, there is some dialectal variation within this language, including some phonological innovations discussed in Clark (2006), who splits Tolomako into three main dialects. Meyerhoff (2009) reports that the speakers of the Matantas variety consider their language to be distinct from the Tolomako varieties spoken further to the west.

Finally, a comment on the classification of these varieties is in order. Above, the following varieties were identified as corresponding to previously labelled languages (with Glottocodes):

- 01 - Nokuku (noku1237),
- 05 - Tasmate (tasm1246),
- 08+09 - Kula/Wusi (wusi1237),
- 10 - Tie (a.k.a Navut) (navu1237),
- 11 - Tiale (a.k.a Malmariv) (tial1239),
- 16 - Piamatsina (piam1242),
- 17 - Tolomako (tolo1255).

Adopting Glottolog's (Hammarström et al. 2026) classification of the languages of Santo, which is based on a reconciliation of Lynch & Crowley (2001: 56-57), Clark (2009a), Tryon (2010), and Lynch (2020b), all of these varieties belong to the West Santo clade (Austronesian > Malayo-Polynesian > Oceanic > North and Central Vanuatu > Espiritu Santo > West Santo) with further subgrouping as in Figure 1. Figure 1 also shows that eight of the varieties that can reliably be paired with Glottolog languages represent four of the seven primary branches of West Santo. The classification of the remaining nine of the 17 varieties is unclear. Based on their geographic distribution, they are very likely to be also part of the West Santo branch (its sister clade, Nuclear Santo, includes languages spoken along the eastern coast and in the southeastern corner of Santo).

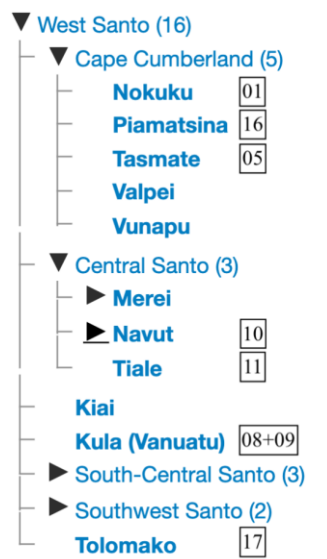


Figure 1. Glottolog 5.3's classification of West Santo languages. The numbers in black boxes indicate the varieties that very likely correspond to these languages.

3. Data and methods

The historical-comparative phonological analysis that follows is based on cognate coded lexical data collected by the author in the 17 locations in 2022, 2023, and 2025. The lexical data include basic and other culturally salient vocabulary. Up to over 700 concepts were elicited for each variety. This study uses data for over 400 concepts for which the data are rich enough. Some concepts that are prone to borrowing or compounding (especially those involving concepts already on the list) have been excluded.

Cognate sets were matched with available reconstructions for Proto North-Central-Vanuatu (PNCV) from Clark (2009a) and Proto Oceanic (POc) from Ross et al. (1998, 2007, 2008, 2011, 2016, 2023).

The supplementary materials include the entire dataset, the subset of it that is used for this study, spreadsheets with cognate sets organised according to reflexes of PNCV sounds, and a detailed description of the methods used in data collection, annotation and organisation, and cognate coding.

4. Vowels

This section discusses the reflexes of PNCV vowels in the 17 varieties. PNCV had the five vowels *a, *e, *i, *o, *u, largely reflecting the same POc vowels (Clark 2009a: 15-16). In the 17 varieties, these vowels also mostly remained unchanged, with some exceptions. The most notable change is the presence of a sixth, likely phonemic, vowel in variety 01. Furthermore, final vowels were lost in many cases in a number of varieties, which is not uncommon in NCV languages (cf. Clark

2009a), perhaps most consistently in variety 01, likely reflecting a relatively recent change (Clark 2009a: 27).

4.1 PNCV *a

PNCV *a remained /a/ in all varieties in all positions (Table 1), except in 01_Nokuku_Tanokuk, where it was raised before a high vowel (*i,*u) in a following syllable (see Clark 2009a: 27 for an earlier discussion of this matter). Examples appear in Tables 2a,2b. The change was most commonly to a back rounded open vowel, whose phonetic value appears to be most commonly [ɔ], as in *kn* ‘tree’ < *kayu and *tms* ‘sea’ < *tasi, but also to [o,e] before *u (*toun* ‘year’ < *ta?u-ni, *meur* ‘live’ < *ma?uri) and [e] before *i (*keriβ* ‘rat’ < *karivi). In the last two examples in Table 2b, *a did not change.

It appears that the process worked on instances of *a in all (adjacent) syllables preceding the triggering *u, as evidenced by *sonoβul* ‘ten’ < *sagavulu and *pernβ* ‘long’ < *baravu. There are no such examples with *i as a trigger in the data set. It remains an open question whether these various reflexes are conditioned.

language	reflexes	BANYAN	WHITE-EYE (Zosterops)	FIVE	DIE	PLANT(v.)	TWO
POC		*ba[k,g]a		*lima	*mate		*rua
PNCV		*baqa	*laka=laka)	*lima-	*mate	*lavo	*rua
01-NKK ⁵	a, [ɔe]/_*[u,i]	pa?	laklak	lima	mat	laβo	rua
02-PNR	a	pa?	lahalah	lima	mate	laβoa	rua
03-PTW	a	pa?a	lahalaha	lima	ma:te	lawoe	rua
04-WNV	a	paka	layalaya	-lima	mate	laβoja	-rua
05-TSM	a	napaka	la?ala?a	lima	mate	laβoja	rua
06-WNR	a	paka	lahalaha	lima	mate	laβo?	rua
07-ELI	a	paka	la?ala?a	lina	na:ti	laβo?e	rua
08-KRP	a	paka	la?ala?a	lima	mat	laβo	rua
09-WSI	a	napaka	la?ala?a	lima	ma:ti	la:βo	rua
10-LND	a	paka	lalao	-lima	ma:te	la:βo	-rua
11-TIA	a	paka	lalao	-lima	tenmate	-laβo	-rua
12-KTN	a	paka	lalao	-lima	tanmate	-laβlaβo	-rua
13-PNS	a	paka	la?ala?a	lina	na:t	laβo	rua
14-WNS	a	te paka	e lalao	ta-lima	ma:te	-laβo	-rua
15-PTL	a	napa?a	lahalaha	lima	mate	laβoe	rua
16-PTS	a	napaka	lahalaha	lima	ma:te	laβoe	rua
17-TLM	a	taka	nalaklak	lina	nate	laβoj	rua

Table 1. Reflexes of PNCV *a without *u or *i in a following syllable.

language	reflexes	WOOD	YEAR	LIVE	LONG	TEN
----------	----------	------	------	------	------	-----

⁵ For brevity in the tables hereinafter I refer to the languages by three-letter codes. These do not necessarily correspond to the respective ISO codes, where such exist.

POC		*kayu	*taqun	*maqurip	*b ^[w] arapu	*sa ŋa puluq
PNCV		*kayu	*taʔu-ni	*maʔuri	*baravu	*sagavulu
01-NKK	a, [veo]/_*[u,i]	kɔ	toun	meur	perɔβ	sonoβul
02-PNR	a	haw	taun	maur	paraβ	sanaβul
03-PTW	a	haw	taun	maur	paraβ	senabul
04-WNV	a	h̄aw	taun	mauri	paraβ	aβla-tea
05-TSM	a	ʔaw	taon	mauri	paraβu	sanaβul
06-WNR	a	ʔaw	taon	mauri	paraβu	sanaβulu
07-ELI	a	ʔaw	taon	nauri	paraβu	sanaβulu
08-KRP	a	ʔaw	taoni	maur	paraβu	hanaβul
09-WSI	a	ʔaw	taoni	mauri	paraβu	hanaβulu
10-LND	a	wuj-ao	taon	mauri	paraβ	sanaβul
11-TIA	a	wi-ao	taun	tenmauri	taparaβ	sanaβul
12-KTN	a	wi-ao	taun	tanmauri	ta-paraβũ	-sanaβul
13-PNS	a	ʔaw	taun	nauri	parawi	sanaβul
14-WNS	a	wi-ao	tauni	mauri	paraβo	-sanaβulu
15-PTL	a	haw	taun	mauri	paraβu	haβla-tea
16-PTS	a	haw	taun	mauri	paraβu	senabulu
17-TLM	a	aɣao	taun	nauri	taraβ	xaβla-tea

Table 2a. Reflexes of PNCV *a before *u and *i in the following syllable (continues as Table 2b).

language	reflexes	SEA	FLYING FOX	RAT	BIRD	FOUR
POC		*tasik			*manuk	*pat
PNCV		*tasi	*qarai	*karivi	*manu-ku	*vati
01-NKK	a, [veo]/_*[u,i]	tds	ɔrɔ	keriβ	man	βati
02-PNR	a	tas	arae	hariβ	mano	vati
03-PTW	a	tas	karae	hariβ	mano	vat
04-WNV	a	ta:si	karae	ʔariβ	manu	-βat
05-TSM	a	tasi	karaj	ʔariβi	mano	βati
06-WNR	a	ta:si	karaj	ʔariβi	mano	βati
07-ELI	a	tas	karaj	ariβi	nano	va:ti
08-KRP	a	tahe	karaj	ʔariβ	mano	βati
09-WSI	a	tahi	karaj	ʔariβi	ma:nu	βa:ti
10-LND	a	ta:si	karaj	arivi	manu	-βat
11-TIA	a	ta:si	karae	ariβĩ	manu	-va:t
12-KTN	a	ta:si	karae	ariβi	manu	-va:ti
13-PNS	a	ta:s	karaj	ariβi	na:no	βa:ti
14-WNS	a	ta:si	karae	ariβi	ma:no	-βa:ti
15-PTL	a	tasi	ʔaraj	hariβi	mano	φati
16-PTS	a	ta:si	karaj	hariβi	mano	βa:ti
17-TLM	a	tas	karaj	hariβi	nanu	vati

Table 2b. Reflexes of PNCV *a before *u and *i in the following syllable (cont.).

4.2 PNCV *e

PNCV *e remained [e] in most cases (Table 3). In the last two examples in Table 3, I have highlighted a few cases where *e changed to [a,i] in word-final (unstressed) position (green) or was lost (blue).

language	reflexes	TARO	RIPE	WHO	STONE	DIE
POC			*mena (Clark 2009a)			*mate
PNCV		*bweta	*mena	*sei	*sule (Clark 2009b – Proto Santo)	*mate
01-NKK	e	p ^w et	men	se cenɔ	sul	mat
02-PNR	e	p ^w eta	mena	sehe ej	sule	mate
03-PTW	e	p ^w eta	mena	sej	sule	ma:te
04-WNV	e	p ^w eta	tamena	sej	sule	mate
05-TSM	e	p ^w eta	mena	sej	sule	mate
06-WNR	e	peta	mena	isei	sule	mate
07-ELI	e	p ^w ete	mene	sej	sula	na:ti
08-KRP	e	p ^w ete	mene	hej	hula	mat
09-WSI	e	p ^w ete	mene	hej	hula	ma:ti
10-LND	e	peta	mena		su:le	ma:te
11-TIA	e	peta		sej	sule	tenmate
12-KTN	e	peta		isej	esule	tanmate
13-PNS	e	p ^w ete	mene	sej	sula	na:t
14-WNS	e	peta	mena		sule	ma:te
15-PTL	e	peta	mena	sej	sule	mate
16-PTS	e	p ^w eta	mena	sej	sule	ma:te
17-TLM	e		nen	sej	sule	nate

Table 3. Reflexes of PNCV *e.

4.3 PNCV *i

PNCV *i remained [i] in most cases (Table 4). I have highlighted a few examples where *i was lost word-finally (blue), or changed to [e] (yellow), especially in a number of varieties in the concept NAME where an early *a > *e change, metathesis and assimilation may have played a role.

language	reflexes	KINGFISHER (Halcyon chloris)	FIVE	SEVEN	RAT	DAY	SEA	NAME
POC		*[s,j]iko	*lima	*pitu		*rani	*tasik	*[q]aca-; *i[c,s]a-
PNCV		*siko	*lima-	*bitu	*karivi	*rani	*tasi	*kisa > *asa
01-NKK	i	sik	lima	pitu	keriβ	rɔn	tɔs	kise-m
02-PNR	i	siho	lima	pitu	hariβ	ran	tas	his
03-PTW	i	siho	lima	pitu	hariβ	ran	tas	hise
04-WNV	i	siyo	mo-lima		ʔariβ	moran	ta:si	ise-
05-TSM	i	siʔo	lima	pitu	ʔariβi	ran	tasi	ise
06-WNR	i	nasifo	lima		ʔariβi	ran	ta:si	yise
07-ELI	i	siʔo	lina		ariβi	rani	tas	ese-
08-KRP	i	hiʔo	lima		ʔariβ	rani	tahe	heʔe-

09-WSI	i	hiʔo	lima		ʔariβi	rani	tahi	ehe
10-LND	i	sio	mo-lima		arivi	ra:ni	ta:si	e:se-
11-TIA	i	sio	mo-lima		ariβi	rani	ta:si	esa-
12-KTN	i	sio	mo-lima		ariβi	ta-ran	ta:si	ese-
13-PNS	i	siʔo	lina		ariβi	ran	ta:s	ese-
14-WNS	i	e sio	ta-lima		ariβi	ta-rani	ta:si	ese-
15-PTL	i	nasiho	lima		hariβi	ran	tasi	hise-
16-PTS	i	siho	lima	pitu	hariβi	ran	ta:si	hisej-
17-TLM	i	siyo	lina		hariβi	ran	tas	hise-

Table 4. Reflexes of PNCV *i.

4.4 PNCV *o

PNCV *o remained [o] in most cases (Table 5). Cases where *o was lost word-finally in the concept SWIM are highlighted in blue. In the cases highlighted in green in HEAR, *o was likely lost, and [e] is a reflex of the POC transitive suffix *i (Lynch et al. 2002: 80), which is also reflected in other varieties that did not lose the stem-final *o.

language	reflexes	THREE	CHICKEN, FOWL	BUY	ORANGE (CITRUS)	PLANT (v)	SWIM	HEAR
POC		*tolu	*toqa	*poli (v)	*molis		*lo- loso[p]	*[l,r]onjon; *ronjor-
PNCV		*tolu	*toʔa	*voli (*vuli)	*moli	*lavo	*loso-vi	*rogo
01-NKK	o	tolu	toa	woli	mol	laβo	lolos	rono
02-PNR	o	tolu	toa	woli	molu	laβoa	lolos	rono
03-PTW	o	tolū	toa	wolia	mol pej	lawoe	loloso	ronjoe
04-WNV	o	-tolu	toa	wulia	moli	laβoja	loloso	ronoja
05-TSM	o	tolu	toa	wolia	moli	laβoja	loloso	ronoja
06-WNR	o	tolu	toa	woli	moli	laβoʔ		rone
07-ELI	o	tolu	to:	vol	moli	laβoʔe	loloso	rone
08-KRP	o	tolu	toa	woli	moli	laβo	lololoho	rone
09-WSI	o	tolu	toa	woli	moli	la:βo	loloho	rone
10-LND	o	-tol	toa	wol=wol	moli	la:βo	loloso	pa:βrono
11-TIA	o	-tol	toa	-βoli	moli	-laβo	ko-loloso	paβrono
12-KTN	o	-tolu	toa	-βoli	moli	- laβlaβo	loloso	-rono
13-PNS	o	tolu	toa	βoli	moli	laβo	lolos	rone
14-WNS	o	-tolu	toa	-βoli	moli	-laβo	-loloso	-rono
15-PTL	o	tolu	toa	wolia	moli	laβoe	loloso	rohoj
16-PTS	o	tolu	toa	βolia	namol	laβoe	loloso	ronjoe
17-TLM	o	tolu	toa	βolia	mol	laβoj	lolos	royoj

Table 5. Reflexes of PNCV *o.

4.5 PNCV *u

PNCV *u remained [u] in all positions (Table 6). Word-final *u was lost in varieties 01,02 in LOUSE, but not in THREE.

language	reflexes	RAIN	STONE	HILL	LIVE	FEATHER	LOUSE	THREE
POC		*qusan			*maqurip	*pulu[-pulu]	*kutu	*tolu
PNCV		*ʔusa	*sule (Clark 2009b–Proto Santo)	*vuti	*maʔuri	*vulu	*kutu	*tolu
01-NKK	u	us	sul	wut	meur	wulu-	kut	tolu
02-PNR	u	usa	sule	wut	maur	wulu	hu:t	tolu
03-PTW	u	usa	sule	wutĩ	maur	βulũ	hutũ	tolũ
04-WNV	u	usa	sule	naβuti	mauri	βulu-	ʔutu	-tolu
05-TSM	u	usa	sule	wuti	mauri	βulu-	ʔutu	tolu
06-WNR	u	u:sa	sule	patiwuti	mauri	wulu-	xutu	tolu
07-ELI	u	usa	sula	p ^w atwuti	nauri	wulu-	ʔutu	tolu
08-KRP	u	uha	hula	p ^w atut	maur	βulu-	ʔutu	tolu
09-WSI	u	uha	hula	p ^w atwuti	mauri	wulu-	u:tu	tolu
10-LND	u	usa	su:le	patwut	mauri	fulu-	utu	-tol
11-TIA	u	usa	sule	patwuti	tenmauri	βulu-	utu-	-tol
12-KTN	u	usa	esule	patwut	tanmauri	wulu-	utu	-tolu
13-PNS	u	usa	sula	p ^w atut	nauri	βulu-	utu	tolu
14-WNS	u	eusa	sule	patwuti	mauri	βulu-	e-utu	-tolu
15-PTL	u	usa	sule	wuti	mauri	wulu-	hutu	tolu
16-PTS	u	usa	sule	wuti	mauri	βulu-	hutu	tolu
17-TLM	u	usa	sule	βut	nauri	βulu-	xutu	tolu

Table 6. Reflexes of PNCV *u.

5. Consonants

This section offers a historical-comparative phonological analysis of the consonants in the 17 varieties. I identify regular sound changes and discuss some irregularities. Section 5.8 summarises the evolution of the varieties' consonant inventories and discusses consonant sound correspondences among them.

The presentation is organised around PNCV phonemes, in the order in which they appear in Table 7, which presents the PNCV phoneme inventory as posited by Clark (2009a: 10ff).

=====

		labial	labial labiovelarised	coronal	velar	glottal
plosives	prenasalised	*b	*b ^w	*d	*q	
	plain			*t	*k	*ʔ
affricate				*z		
nasals		*m	*m ^w	*n	*g	
fricatives		*v	*v ^w	*s		
lateral				*l		
rhotics				*r	*R	
approximants			*w	*y		

Table 7: The PNCV consonant inventory (Clark 2009a: 10ff). The symbols are as in the source.

All PNCV reconstructions are from Clark (2009a) and all POC reconstructions are from Ross et al. (1998, 2007, 2008, 2011, 2016, 2023), unless otherwise stated. Sound correspondences between POC and PNCV are mostly based on Clark (2009a: 16).

The following systemic developments in the consonant inventories of the 17 varieties stand out:

- in all varieties, the prenasalisation contrast has been lost (with some minor exceptions, the most regular one being the retention of a prenasalised trill in 16_Piamatsina_Pwelmau), with the contemporary varieties having a single plosives series;
- contrastive labiovelarisation as secondary articulation has been completely lost in some varieties and retained to various degrees in the nasals and/or plosives in others; it has been lost in all varieties in the fricatives.

These developments have also been pointed out by Clark (2009b) for the relevant varieties also included in that study. These and other observations are discussed in detail in the relevant subsections.

5.1 The PNCV plosives

5.1.1 PNCV *b

PNCV prenasalised *b, which is inherited from the same POC prenasalised *b, regularly developed into a plain /p/ in all varieties with the following exceptions (Tables 8a, 8b):

- In variety 17, *b yielded [t] before non-rounded vowels, which, as discussed in Clark (2009b), is the result of an earlier bilabial (*b in this case) first changing to the cross-linguistically rare linguolabial ([ⁿd]) or [t̚] in this case, depending on when loss of

prenasalisation occurred relative to the linguolabial shift), and later merging with the coronal consonants (/t/ in this case). In Clark's (2009b) terms, this is the Type II linguolabial shift (see also Lynch 2020b on this topic).

- In 14_Winsao_Tie *b is reflected as [p] or [mp], see last example in Table 8b and supplementary material. The reflex [mp] is likely a complex segment – a prenasalised bilabial plosive, which lacks clearly audible voicing during the complete airflow obstruction. Hence, I note this segment as <^mp>. In this variety, stress appears to consistently fall on the penultimate syllable and the prenasalised variant appears to be restricted to before non-stressed vowels. See also §5.1.3 and §5.1.4.

language	reflexes	BANYAN	LONG	KNOW	SMELL
POC		*ba[k,g]a	*b ^[w] arapu		*bo-na
PNCV		*baqa	*baravu	*bi(ni)si	*boni (*buni)
01-NKK	p	paʔ	perɔβ		ponpon
02-PNR	p	paʔ	paraβ		poni
03-PTW	p	paʔa	paraβ		ponpon
04-WNV	p	paka	paraβ		ponina
05-TSM	p	napaka	paraβu	pinis	ponpon
06-WNR	p	paka	paraβu	pinis	ponponis
07-ELI	p	paka	paraβu		pono
08-KRP	p	paka	paraβu		ponopono
09-WSI	p	napaka	paraβu	pinih	pon ^d rone
10-LND	p	paka	paraβ	pipinisi	ponaga
11-TIA	p	paka	taparaβ		ponana
12-KTN	p	paka	ta-paraβū		ponana
13-PNS	p	paka	parawi		ponoina
14-WNS	p, ^m p [unstressed V]	te paka	paraβo	pipinisi	ponana
15-PTL	p	napaʔa	paraβu	pinis	punina
16-PTS	p	napaka	paraβu	pisi	ponpon
17-TLM	t, p/ ^m p *[o,u]	taka	taraβ	tinis	pon

Table 8a. Reflexes of PNCV *b (continues as Table 8b).

language	reflexes	HIDE	MOON	WING	FIRE
POC		*buni, *puni		*kaba[-]	*[k,g]abu
PNCV			*kabatia (*kabatio)	*kaba-u	*kabu
01-NKK	p	susurpun			
02-PNR	p	puniā	hapatia		hap
03-PTW	p	punia	hapatia		hapu
04-WNV	p	punenia	apatia		a:pu
05-TSM	p	punia	apatia		ʔapu
06-WNR	p	punen			a:pu
07-ELI	p	punen		apau-	
08-KRP	p	punen		apau-	
09-WSI	p	puneni		apau-	
10-LND	p	tani punaj		pa=pau-	a:pu
11-TIA	p	tani punaj		pa=pau-	a:pu

12-KTN	p	tani punaj		pa=pau-	a:pu
13-PNS	p	tʃenpunen		apau-	
14-WNS	p, ^m p [unstressed V]	tapunae		pa=pau-	a: ^m po
15-PTL	p	punia	hapatia		hapu
16-PTS	p	lopunia	hapatia		hapu
17-TLM	t, p/ *[o,u]	punia	hatatia		xapu

Table 8b. Reflexes of PNCV *b (cont.).

5.1.2 PNCV *b^w

The southernmost varieties and variety 17 consistently reflect prenasalised *b^w as a plain [p]. In the rest of the varieties labiovelarisation has been retained, at least in some examples (see Table 9 and Map 2). It needs to be pointed out that the presence or absence of labiovelarisation may depend on free allophony or intra- and inter-speaker variation. Clark (2009a: 11) reports such crossover between labiovelarised and plain plosives in other Vanuatu languages. In four of the 17 varieties – 05, 07, 13, 16 – [p^w] is attested consistently. As labiovelarised sounds normally do not occur before rounded vowels, the [p] reflexes in 01_Nokuku_Tanokuk *pɔlkɔ- taleman* ‘father-in-law’ and *potu-* ‘head’ can be explained by a proto *a changing to a rounded vowel (cf. Clark 2009a: 27), thus neutralizing labiovelarisation.

language	reflexes	LIVER	TARO	FATHER-IN-LAW	KNEE	HEAD	OWL	HALF
POC					*b ^w au- (PSOc) *p ^w atu[ka]	*b ^w atu[k]	*b ^w eka	
PNCV		*mwabwe	*bweta	*bwalika	*bwau	*bwatu	*bwek[a,e] (TR ⁶)	*sobwe
01-NKK	p ^w /(p)	map ^w e-	p ^w et	pɔlkɔ- taleman		potu-	p ^w e	sop ^w en
02-PNR	p ^w /(p)	map ^w e	p ^w eta	palha-		patu-	p ^w e?	sop ^w en
03-PTW	p ^w /(p)	map ^w e-	p ^w eta	p ^w aliha	pau?u	p ^w atu-	p ^w eka	
04-WNV	p ^w /(p)	map ^w e-	p ^w eta	p ^w alia-	paum		p ^w eka	sop ^w ena
05-TSM	p ^w	map ^w e-	p ^w eta	p ^w aliya	p ^w auna		p ^w eka	sop ^w ena
06-WNR	p/p ^w	ma:pe	peta	palixa-	p ^w au		p ^w ece	sop ^w ena
07-ELI	p ^w	majm ^w ap ^w e	p ^w ete	pali?a / p ^w al?a-	p ^w auna	p ^w atu-	p ^w ece	wasp ^w e ne
08-KRP	p ^w /(p)	m ^w ap ^w e-	p ^w ete	pali?a / p ^w a?la-	paum	p ^w atu-	p ^w eke	hup ^w ene
09-WSI	p ^w /(p)	m ^w ap ^w e-	p ^w ete	palia	pauna	patu-	p ^w elulu	hup ^w ene
10-LND	p	ma:pe-	peta	pali	pauka	patu-		vaspena
11-TIA	p	mape-	peta	palia-	pauka	patu-		
12-KTN	p	mape-	peta		pauka	patu-		
13-PNS	p ^w	m ^w ap ^w e-	p ^w ete	p ^w al?a-	p ^w aum	p ^w atu-	p ^w ece	
14-WNS	p	mape-	peta	palia- tauku	pau ^u ka	pa:tu-		supena

⁶ Here and elsewhere, TR means that the reconstruction is my own.

15-PTL	p	majmape	peta	paliha-	paum		peʔa	sopena
16-PTS	p ^w	m ^w ap ^w e-	p ^w eta	p ^w alha-	p ^w aum	p ^w atu-	p ^w eka	sop ^w ena
17-TLM	p	nape-		paliya tatʃua	paum			

Table 9. Reflexes of PNCV *b^w.Map 2. Reflexes of PNCV *b^w. Base map: Google Maps; annotations by the author (here and in the following maps).

5.1.3 PNCV *d

PNCV *d stands for a prenasalised coronal plosive, which is a regular reflex of the same or very similar POC *d as well as POC *dr, a prenasalised coronal trill, where trilling starts from a closed position resulting in an audible plosive-like release following the nasal phase. This complex segment, also known as the “NDR-sound”, a cover term for an array of similar phonetic variants (cf. Rangelov 2023, 2024), is found in a couple of dozen languages on Santo, Malekula and Efate in Vanuatu, as well as other Oceanic and Austronesian languages. While Clark (2009a) does not include *dr in the phoneme inventory of PNCV, he notes that a distinction between *d and *dr persisted in some NCV languages (cf. Clark 2009a: 11).

Tables 10a, 10b and Map 3 show the reflexes of PNCV *d. An NDR-sound is a regular reflex of *d only in 16_Piamatsina_Pwelmau. In the rest of the varieties, it is reflected as either (1) glottal stop or zero in apparently free allophony, probably related to the low perceptual salience of glottal stops, especially at word boundaries, (2) the velar stop [k], or (3) an affricate ([tʃ]) or, less commonly, as what I have denoted as [ts̥], [t̥] or [t̥ʃ] in an attempt to capture phonetic detail, i.e. an

affricate or, in a couple of examples, a plosive, articulated somewhat further forward than a $[\widehat{tj}]$.⁷ The pathway from a $[^nd]/[^ndr]$ to $[k]/[ʔ]$ likely involved intermediate stages. There is clear evidence for loss of prenasalisation in all consonants and for velarisation of the coronal plosive that resulted in a merger of $*q$ $[^ng]$ and $*d$ into a plain velar plosive (see §5.1.4 and §5.8 for a discussion of $*q$ and the relative timing of the changes). In the varieties where we see affricates, instead of velarisation, the coronal plosive appears to have undergone palatalisation; alternatively, starting from an NDR-sound, after loss of prenasalisation, what remains is a plosive release and a rhotic, of which the latter, as per cross-linguistic tendencies, may surface as a non-trilling allophone, including a fricative (cf. Ladefoged and Maddieson 1996: 217), i.e. $*_{ndr} > *_{dr}/r > tʃ/ts$.

There are some exceptions to the general patterns, highlighted in red in Tables 10a, 10b:

- In 03_Petawata_Hurahi $*d$ is irregularly reflected as $[k]$ in *koko* ‘cloud’, possibly due to borrowing from the neighbouring varieties to the south. In the same variety, the occurrence of an NDR-like sound in $^{nd}lo.^{nd}lo$ ‘night’ is unusual. This is the only occurrence of a similar sound in this variety, which otherwise does not have any prenasalised segments. I transcribe this using the retroflex approximant $ɭ$ to note phonetic detail; it is not unusual for the NDR-sound to have retroflex and approximant variants. The example’s clear similarity to 16_Piamatsina_Pwelma’s $^{nd}ro.^{nd}ro$ ‘night’, and the reported close historical ties between these communities (§2), suggest that this form may have been borrowed into 03 from 16. We can also speculate that it may be a relic of an earlier, more regular NDR-sound in 03, retained in this item because of the relatively high frequency of the concept NIGHT and its participation in formulaic language with high salience, such as the greeting $^{nd}lo.^{nd}lo$ *tahalel* ‘good night’.
- In 05_Tasmate_Oa, $*d$ is irregularly reflected as a glottal stop in $ʔoa$ ‘no’, as well as in *taʔua* ‘man’ and *ʔalu* ‘worm’ (see additional data in the supplementary material for the latter two examples). It is unclear what the source of these irregularities is. There may be idiolectal effects or, the relatively high salience and token frequency of the concepts, at least the first two, may be relevant. $ʔoa$ is also the form that is the endonym of this variety.
- In 10_Linduri_Tie and 14_Winsao_Tie we see *iʔje* ‘we (1PL_INCL)’ involving a prenasalised voiced plosive (palatalisation is likely due to the following front vowel) instead of the expected plain (normally voiceless) $[c]$. As suggested by a reviewer, in this case, the prenasalised plosives may reflect a longer proto-form $*(i)gida$, which

⁷ These less common reflexes also appear in the data as reflexes of PNCV $*z$, which has an affricate reflex in all varieties (cf. §5.2). These less common reflexes are restricted to varieties 07-12. Of these, varieties 07,08,09 appear to have merged PNCV $*d$ with PNCV $*z$. There does not seem to be a clear pattern suggesting a $[\widehat{tj}] \sim [\widehat{ts}]$ distinction as a retention of the $*d \sim *z$ distinction. Speakers of 08_Kerepua_Tuijo, however, have reported self-awareness of two distinct articulations for their affricates, which they prefer to distinguish in spelling too as $\langle d \rangle$ vs $\langle j \rangle$. However, their spelling preferences do not consistently pattern with my phonetic perceptions or with the reflexes of PNCV $*d$ and $*z$. This may suggest that the merger of $*d$ and $*z$ in this variety, and possibly others, may be still ongoing, perhaps in its final stages. In any case, this issue deserves more attention and perhaps more detailed phonetic and sociophonetic investigations.

would produce the expected form **inike, and after loss of the middle vowel, the resulting sequence of homorganic nasal + plosive surfaces as a prenasalized plosive. The occurrence of [ɲ] here is also somewhat in line with other plosives in these varieties, cf. discussion on ^mp in §5.1.1, and on ^ɲk/ɲg in §5.1.4, the latter, as here, clearly restricted to grammatical morphemes, which explains, at least partially, the persistence of these unusual reflexes – such morphosyntactic conditions may support the persistence of otherwise marginal sounds in a phonological system (Evans 2022, Rangelov 2026b: 7).

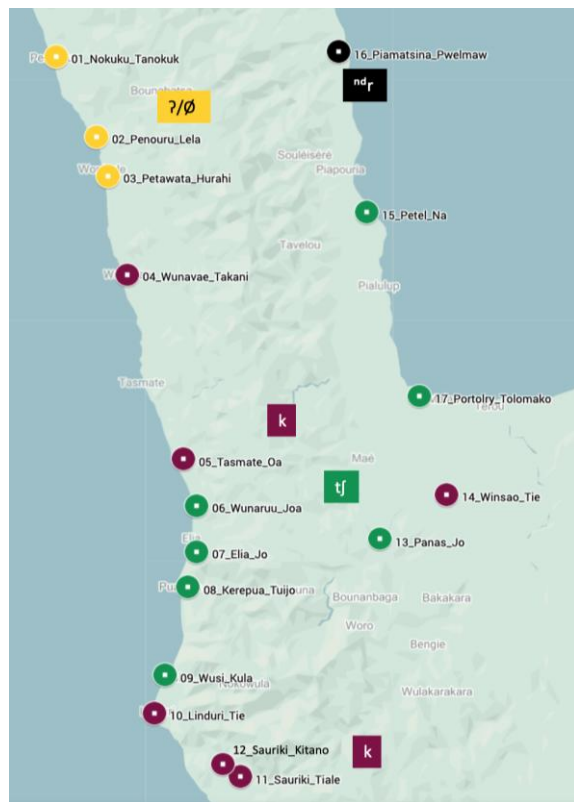
language	reflexes	EAR	NECK	YAM	WE_INCL	THINK	KNOW
POC		*taliŋa		*damu (PEOc)		*drodrom, *drom-i	*drodrom, *drom-i
PNCV		*daliga	*daleʔo	*damu	*(k)ida	*do=dom-i	*do=dom-i
01-NKK	ʔ/Ø	ɳnla-	alo-	ɳm	riʔe		
02-PNR	ʔ/Ø	anla-	alo-	am	a-tol		
03-PTW	ʔ/Ø	anla-	ʔalo	am	a-tolū		
04-WNV	k	kalna-	kalu-	kam	ice-mara	komkom	komkom
05-TSM	k/(ʔ)	kalna-	kalo-	kam	ice	komkom	
06-WNR	tʃ	tʃalina-	tʃalo-	tʃam	itʃa	tʃomtʃom	
07-ELI	tʃ	tʃalna-	tʃolo-		itʃe	tʃojtʃon	
08-KRP	tʃ	tʃalna-	tʃolo-		itʃe	tʃom-im	
09-WSI	tʃ	tʃalna-	tʃolo-		itʃe	tomtom	
10-LND	k	kalna-	kolo-	kam	i ^ɲ je	komkomi	
11-TIA	k	kalna-	kolo-	kam	ice		komkomi
12-KTN	k	kalna-	kolo-	kam	ika wuwut		komkomi-
13-PNS	tʃ	tʃalna-	tʃolo-		itʃe	tʃomi-na	
14-WNS	k	kalna-	kolo-	kam	i ^ɲ je	komkomi- ^ɲ ga	
15-PTL	tʃ	tʃalha-	tʃalo-	tʃam	hitʃe		
16-PTS	nd r	nd ralŋa-	nd ralu-	nd ram	ni nd re		
17-TLM	tʃ	tʃaliya-	tʃalo-		itʃe		

Table 10a. Reflexes of PNCV *d (continues as Table 10b).

language	reflexes	CLOUD	BLOOD	TOOTH	TOOTH	NIGHT	NO
POc		*rodo(ŋ), *rodrom (Clark 2009a)	*draraq	*ɲado-		?*rodrom	
PNCV		*dodo	*daRa	*gado	*kudu (Clark 2009b)	*dodo (TR)	*doa (TR)
POC	ʔ/Ø	oʔ				ʔoʔor	oa
PNCV	ʔ/Ø	oʔ				oʔor	ʔoa
01-NKK	ʔ/Ø	koko				n ^d ɔo:n ^d ɔo	
02-PNR	k	kokora	kai-				
03-PTW	k/(ʔ)	kokora					ʔoa
04-WNV	tʃ	tʃotʃora	tʃae		utʃu		tʃoaʔ
05-TSM	tʃ	tʃotʃo	tʃaʔa		utʃu-		tʃoʔ

06-WNR	tʃ	tʃotʃo	tʃae-		utʃu-		tʃoteh
07-ELI	tʃ	tʃotʃo	tʃae		u:tʃu-		
08-KRP	k	ko:ko	kae-	ako-			
09-WSI	k		kae-	ako-			
10-LND	k		kae-	a:ko-			
11-TIA	tʃ	tʃotʃo	tʃai-		utʃu		
12-KTN	k		kae-	eako-			
13-PNS	tʃ		tʃae-				tʃoa
14-WNS	nd _r					nd _r o nd ro	nd _r oa
15-PTL	tʃ		tʃaj		utʃu-		motʃoa

Table 10b. Reflexes of PNCV *d (cont.)



Map 3. Reflexes of PNCV *d.

It is also worth noting that an NDR-like pronunciation has been attested in varieties other than 16_Piamatsina_Pwelmau. Besides in the 03_Petawata_Hurahi case described above (*n^dʔo:n^dʔo* ‘night’), NDR may also emerge when underlying /nr/ sequences across syllable (and morpheme) boundaries are pronounced with an epenthetic [d], e.g. in 09_Wusi_Kula *pon^drone* ‘smell’ < PNCV *boni ‘smell’ + *rogo ‘hear, smell, feel’⁸; in 17_Portolry_Tolomako, *kan^drua*

⁸ It is also possible that *pon^drone* is a complex predicate (serial verb construction or similar) where the two stems (*pon* and *rone*) can be considered different words rather than one compound form. More detailed morphosyntactic data and analysis would be necessary to determine wordhood here. Regardless, the result would be the same or very similar for the issue at hand here.

‘you two (2DU)’ < PNCV *qamuyu ‘you (pl.)’ + *rua ‘two’; in 13_Panas_Jo, *tan^drum* ‘chief’s house, nakamal’, where the first syllable is likely a reflex of PNCV *tano ‘earth, ground’; in 05_Tasmate_Oa, *ron^drei* ‘dream’, which likely contains a reflex of PNCV *rogo ‘hear, smell, feel’ (*g > n in this variety, cf. §5.3.4). I have annotated these, as here, with a superscript <^d>, rather than as <ndr> to reflect the underlying structure and etymology and that the NDR-sound is not part of the phonological systems of these varieties. The alternative hypothesis that [d] is an underlying segment of its own is unsupported not just etymologically but also by these varieties’ syllable structure, which is (C)V(C), thus constraining CCC clusters. From an articulatory point of view, the epenthetic [d] is likely an artefact of the relative timing of the raising of the velum during the transition between /n/ and /r/ and the start of the trilling (or similar rhotic) gesture in /r/, i.e. there is a short period during which there is a total closure of the airflow, and thus an increase in intraoral pressure, which results in a plosive-like release at the onset of /r/. This articulatory mechanism is very similar to that of the NDR complex segment (see above, cf. Rangelov 2023). From a perceptual and sociophonetic standpoint, this articulation in these varieties might be related to their speakers’ exposure to other Santo varieties with the NDR-sound and/or awareness of NDR as a salient sound. In Tryon’s (1976) data, there is evidence of a likely phonemic NDR-sound in ten Santo varieties, including Piamatsina; see also Rangelov (2023) on the NDR-sound in Farsaf, a language spoken in southern Santo.

5.1.4 PNCV *q

PNCV *q [ʔg] regularly reflects POc *g (which had the same phonetic value). In the 17 varieties, its reflexes (see Table 11 and Map 4) are glottal stop/zero in the three northernmost varieties on the west coast – 01,02,03, as well as in 15. In 15, it is consistently a glottal stop, whereas in the other three varieties the glottal stop and zero reflexes are apparently in free variation, probably due to the low perceptual salience of the glottal stop and/or idiolectal effects. This means that in 01,02,03 *q and *d have largely merged (see also Clark 2009b).

In the rest of the varieties, *q is reflected as [k], which means that the *q~*d merger extends also to varieties 04,05 and 10,11,12,14. The fact that this merger is observed in nine out of the seventeen varieties, suggests that the *q~*d merger may have happened relatively early (see also §5.8). It involved a velarisation of *d before or after loss of prenasalisation across the board, later followed by a further debuccalization (and in some cases glottal stop loss) in varieties 01,02,03. Debuccalisation of *k also happened independently in 15_Petel_Na, where the *q~*d distinction is retained.

There are some exceptions to the general patterns, which are highlighted in red in Table 11:

- In 03_Petawata_Hurahi the most common reflex is glottal stop/zero, with a few examples of [k] too, including *karae* ‘flying fox’, *p^weka* ‘owl’, *kala* ‘lizard’ (see also supplementary material) and the first-person possessive suffix, which appears as *-ku* and *-ʔu* in apparently free variation.

- In 10_Linduri_Tie and 14_Winsao_Tie there are a few examples of prenasalised [ɲk]/[ɲg] where [k] is expected. This is in line with the corresponding reflexes of *d (see §5.1.3) and only observed in grammatical morphemes in the data, cf. discussion on ^mp in §5.1.1.
- Regarding the concept YOU_SG, it appears that some varieties reflect a POc oral-grade *k and others, a POc nasal-grade *g, which is also reconstructed for PNCV, but probably both variants existed. The northernmost varieties on both sides of the mountain chain, namely 01-04 and 15-17 reflect *iko, while the rest reflect *iqo.

language	reflexes	BANYAN	DRINK	DRY	FLYING FOX	YOU_PL	WE_EXCL	1SG. POSS	YOU_SG
POc		*ba[k,g]a	?*[g,k] ^w a[g,ŋ] ^w a	*goru					*iko (Clark 2009a)
PNCV		*baqa	*(^ɲ go=) ^ɲ go (TR)	*qora; *qoru	*qarai	*qamuyu	*qama(m)i	*-qu	*iqo
01-NKK	ʔ/Ø	paʔ	ʔo	or	ɔrɔ	omto / emiu	omoto	-ʔ	nik
02-PNR	ʔ/Ø	paʔ	ʔo	orū	arae	amto	amato	-ʔ	niho
03-PTW	ʔ/Ø/(k)	paʔa	ʔo	oru	karae	amto	atow	-ʔu/ -ku	niho
04-WNV	k	paka	koko	koru	karae	kamiu	kamam	-ku	niʔo
05-TSM	k	napaka	koko		karaj	kamju	kamamu	-ku	niko
06-WNR	k	paka	koko	koru	karaj	kamiu	kanam	-ku	niko
07-ELI	k	paka	koko	koru	karaj	kaniu	kanaj	-ku	niko
08-KRP	k	paka	koko	koru	karaj	kamiu	kamaj	-ku	niko
09-WSI	k	napaka	koko	koru	karaj	kamiu	kamaj	-ku	niko
10-LND	k	paka	koko		karaj	kamiu	kamaw	-ka	i ^ɲ go
11-TIA	k	paka	koko		karae	kamiu	kami	-ka	iko
12-KTN	k	paka	koko:ko		karae	komiu	komao	-ka	iko
13-PNS	k	paka	koko	koru	karaj	kaniu	kanaj	-ku	niko
14-WNS	k/ ^ɲ k/ ^ɲ g	te paka	kukoko	-koro	karae	^ɲ qamiu	i ^ɲ qamau	- ^ɲ ka	i ^ɲ go
15-PTL	ʔ	napaʔa	ʔoʔo		ʔaraj	ʔamiu	ʔamam	-ʔu	niho
16-PTS	k	napaka	koko	koru	karaj	kamiu	kamamo	-ku	niho
17-TLM	k	taka	koko	koru	karaj	kaniu	kanam	-k(u)	niyo

Table 11. Reflexes of PNCV *q.



Map 4. Reflexes of PNCV *q.

5.1.5 PNCV *t

PNCV *t (< POc *t) is reflected as /t/ in all varieties without exceptions. It is a frequent token in the data set and there are many examples in the supplementary materials besides those in Table 12.

language	reflexes	SEA	THREE	FOUR	LOUSE	EGG	EYE
POC		*tasik	*tolu	*pat	*kutu	*qatolur	*mata[-]
PNCV		*tasi	*tolu	*vati	*kutu	*ʔatolu	*mata
01-NKK	t	tos	tolu	βati	kut	otolin	meto- / mat
02-PNR	t	tas	tolu	vati	hu:t	otolu	mata-
03-PTW	t	tas	tolū	vat	hutū	otolū	
04-WNV	t	ta:si	-tolu	-βat	ʔutu	otoli	mata-
05-TSM	t	tasi	tolu	βati	ʔutu	otolina	mata-
06-WNR	t	ta:si	tolu	βati	xutu	tolina	mata-
07-ELI	t	tas	tolu	va:ti	ʔutu	tolina	nata-
08-KRP	t	tahe	tolu	βati	ʔutu	tolina	mata-
09-WSI	t	tahi	tolu	βa:ti	u:tu	tolina	mata
10-LND	t	ta:si	-tol	-βat	utu	tolina	mata-
11-TIA	t	ta:si	-tol	-va:t	utu-	tolina	mata-
12-KTN	t	ta:si	-tolu	-va:ti	utu	tolina	mata-
13-PNS	t	ta:s	tolu	βa:ti	utu	tolina	nata-
14-WNS	t	ta:si	ta-tolu	ta-βa:ti	e-utu	tolina	mata-

15-PTL	t	tasi	tolu	ɸati	hutu	otol	mata-
16-PTS	t	ta:si	tolu	βa:ti	hutu	otoli	mata-
17-TLM	t	tas	tolu	vati	xutu	oyotol	nata-

Table 12. Reflexes of PNCV *t.

5.1.6 PNCV *k

PNCV *k (< POc *k) regularly remained [k] only in 01_Nokuku_Tanokuk. In the rest of the varieties, it developed into either (1) a velar/glottal fricative, (2) a glottal stop/zero, (3) only as zero, or (4) as a combination of those (see Tables 13a, 13b, 13c). Three clusters emerge, cf. Map 5.

Starting from the north, in varieties 02,03 and 15,16 (straddling the mountain chain), the reflex is consistently the glottal fricative [h]. 17_Portolry_Tolomako may also be considered part of this cluster, as, in it, the reflexes of *k are fricatives, either [h] or a voiced or voiceless velar fricative. It is unclear what the distribution of these phones is in this variety. For example, Meyerhoff (2009: 5) reports a velar and glottal fricative as separate phonemes in the Matantas variety, which may be considered part of the same language (cf. §2 above).

In varieties 05,07,08,09,13, the reflexes of *k surface as glottal stop or zero, the latter likely a further development from a glottal stop. The two surface representations appear to be in free variation.

Further south, in varieties 10,11,12,14, *k regularly yields zero.

This leaves two varieties, in which all possible reflexes, except [k], are represented. In 04_Wunavae_Takani, there is a tendency for word-initial *k to be reflected as glottal stop/zero, with a few fricative examples too, while the reflexes in word-medial position tend to be fricative or zero. In 06_Wunaruu_Joa, the situation is reversed: word-initial *k is more often reflected as a fricative and sometimes as a glottal stop/zero, while word-medial *k can be reflected by a glottal stop, as well as zero or a fricative. The relative lack of regularity stands out compared to the high degree of regularity in some of the other varieties.

A few exceptions are highlighted in Tables 13a-13c. The apparent [k] reflexes in 17_Portolry_Tolomako are likely borrowings from Bislama, while *ita* ‘octopus’ may be lacking an initial fricative, as expected, due to perhaps an early loss of the initial syllable containing *k. The unexpected [ɣ] reflex in *p^waliya* ‘father-in-law’ in 05 may be due to borrowing, possibly from 17 (cf. §2).

language	reflexes	WOOD	LOUSE	RAT	NAME	EAT	WHITE-EYE (Zosterops)
POC		*kayu	*kutu		*[q]aca-; *[c,s]a-	*kani	
PNCV		*kayu	*kutu	*karivi	*kisa > *asa	*kani	*laka=laka
01-NKK	k	kɒ	kut	keriβ	kise-	kɒnkɒn	laklak
02-PNR	h/(ɦ)	haw	hu:t	hariβ	his	hanhan	lahalah
03-PTW	h	haw	hutũ	hariβ	hise	han	lahalaha

04-WNV	initial: ʔ/Ø, (fi/ɣ) mid: ɣ/fi, Ø	fiaw	ʔutu	ʔariβ	ise-	ʔanʔan	layalaya
05-TSM	ʔ/(Ø)	ʔaw	ʔutu	ʔariβi	ise	ʔanʔan	laʔalaʔa
06-WNR	initial: ɣ/x/h/, ʔ/Ø mid: ʔ/Ø, fi/ɣ/x	ʔaw	xutu	ʔariβi	ɣise	ɣaneyane	laʔalaʔa
07-ELI	ʔ/Ø	ʔaw	ʔutu	ariβi	ese-	ʔanʔan	laʔalaʔa
08-KRP	ʔ/Ø	ʔaw	ʔutu	ʔariβ	heʔe-	ʔanʔan	laʔalaʔa
09-WSI	ʔ/Ø	ʔaw	u:tu	ʔariβi	ehe	ʔanʔani	laʔalaʔa
10-LND	Ø	wuj-ao	utu	arivi	e:se-	anani	lalao
11-TIA	Ø	wi-ao	utu-	ariβi	esa-	-anan	lalao
12-KTN	Ø	wi-ao	utu	ariβi	ese-	-anan	lalao
13-PNS	ʔ/Ø	ʔaw	utu	ariβi	ese-	ʔanʔan	laʔalaʔa
14-WNS	Ø	wi-ao	e-utu	ariβi	ese-	-anan	e lalao
15-PTL	h	haw	hutu	hariβi	hise-	han	lahalaha
16-PTS	h	haw	hutu	hariβi	hisej-	hani	lahalaha
17-TLM	x/ɣ/h	ayao	xutu	hariβi	hise-	vyan	nalaklak

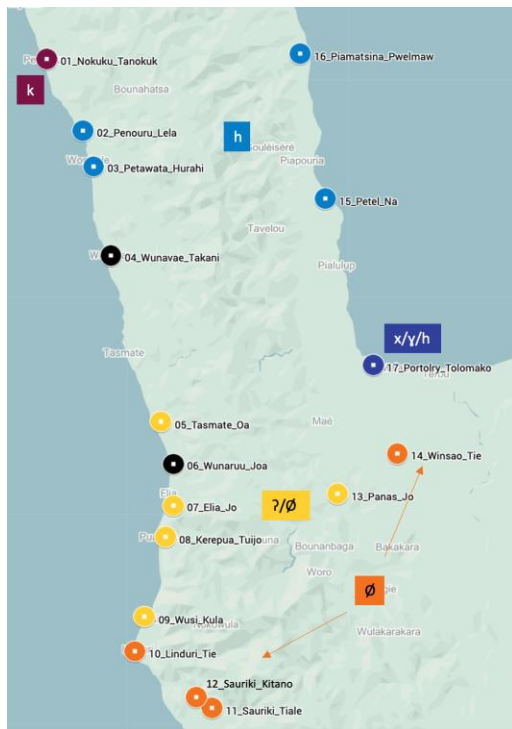
Table 13a. Reflexes of PNCV *k. (continues as Table 13b).

language	reflexes	KING FISHER	CRAB	FIRE	OCTOPUS	STEAL	BITE
POC		*[s,j]iko	*kape	*[k,g]abu	*kurita		*karat-i; *kati[-]
PNCV		*siko	*kave	*kabu	*kuRita	*(bwa=) bwaliko (TR)	*kaRa-ti
01-NKK	k	sik	ka:β		kwit	pɪp ^w elik	kɔt ^w moti
02-PNR	h/(fi)	siho	ha:βe	hap	hwit	pali ^w foe	hat ^w moti
03-PTW	h	siho	haβe	hapu	hwita	pali ^w hoe	
04-WNV	initial: ʔ/Ø, (fi/ɣ) mid: ɣ/fi, Ø	siyo	iyaβe	a:pu	wita	p ^w ap ^w alio	atusia
05-TSM	ʔ/(Ø)	siʔo	ʔaβe	ʔapu	wita	p ^w ap ^w aliʔo	ʔat ^w motia
06-WNR	initial: ɣ/x/h/, ʔ/Ø mid: ʔ/Ø, fi/ɣ/x	nasifo	xaβe	a:pu	nawita	paliʔe	xati tusia
07-ELI	ʔ/Ø	siʔo	ʔaβe		uita	p ^w aliʔe	et tus
08-KRP	ʔ/Ø	hiʔo	aβe		wita	paliʔe	
09-WSI	ʔ/Ø	hiʔo	aβe		uita	paliʔe	
10-LND	Ø	sio	a:ve	a:pu	wita		a:ti
11-TIA	Ø	sio	aβe	a:pu	wita		-atrot
12-KTN	Ø	sio	a:ve	a:pu	wita		-atro:to
13-PNS	ʔ/Ø	siʔo	a:βe		wita	p ^w aliʔe	
14-WNS	Ø	e sio	e a:βe	a: ^m po	nawita		-atroto
15-PTL	h	nasiho	haβe	hapu	nahwita	pali ^w hoe	hat tusia
16-PTS	h	siho	haβe	hapu	hoita	p ^w ali ^w hoe	hat ^w motia
17-TLM	x/ɣ/h	siyo	yaβe	xapu	ita	paliyoj	hatia

Table 13b. Reflexes of PNCV *k. (continues as Table 13c).

language	reflexes	MOSQUITO	SHARK	MALAY APPLE	KAVA	FATHER- IN-LAW	FEAR
POc		*ñamuk	*bakewa	*kapika			*matakut-
PNCV		*namu-ki	*bakewa	*kavika	*maloku	*bwalika	*mataku
01-NKK	k	mok	pace	cevik		pɔlko-	
02-PNR	h/(f)	moh	paheo	haβih		palha-	
03-PTW	h	mohɪ	paheo	haβiha		p ^w aliha	
04-WNV	initial: ʔ/Ø, (f/ɣ) mid: ɣ/f, Ø	mohɪ	payeo	yaβiya	ae	p ^w alia-	
05-TSM	ʔ/(Ø)	moʔi	paʔeo	ʔaβiʔa	maloʔu	p ^w aliya	
06-WNR	initial: ɣ/x/h/, ʔ/Ø mid: ʔ/Ø, f/ɣ/x	moyi	paheo	haβifa	maloxu	palixa-	mata
07-ELI	ʔ/Ø		paʔoe		neloʔu	paliʔa	natau
08-KRP	ʔ/Ø		paʔoe		meloʔu	paliʔa	mataʔo
09-WSI	ʔ/Ø		paʔoe		meloʔu	palia	mataʔo
10-LND	Ø	muj	paeo		malow	pali	mataw
11-TIA	Ø	muj	paeo		malo	palia-	-matao
12-KTN	Ø	muj	paeo		malo		-mamatau
13-PNS	ʔ/Ø		paʔoj		neloʔ	p ^w alʔa-	nataʔo
14-WNS	Ø				malow	palia-	matao
15-PTL	h	mohi		haviha		paliha-	
16-PTS	h	mohi		haβiha		p ^w alha-	
17-TLM	x/ɣ/h	mox		yaβiya	maluku	paliya tatʃua	

Table 13c. Reflexes of PNCV *k. (cont.).



Map 5. Reflexes of PNCV *k. In the two varieties in black, the reflexes vary see Tables 13a, 13b, 13c.

5.1.7 PNCV *ʔ

The PNCV glottal stop *ʔ (< POc *q) is regularly lost in all the varieties, as demonstrated by the examples in Table 14 and other examples in the supplementary materials. In the highlighted example in Table 14 the initial [h] is likely prothetic rather than a reflex of *ʔ. Prothetic [h] is not uncommon in Bislama (likely often due to hypercorrection), in which all recorded speakers are also fluent, which means that this tendency may also extend to the Indigenous languages in some cases.

language	reflexes	YEAR	LIVE	CHICKEN, FOWL	RAIN	CRAYFISH	SMOKE
POc		taqun	*maqurip (v)	*toqa	*qusan	*quraŋ	*qasu
PNCV		*taʔu-ni	*maʔuri	*toʔa	*ʔusa	*ʔura	*ʔasu
POc	Ø	toun	meur	toa	us	ur	osun ɒβ
PNCV	Ø	taun	maur	toa	usa	ura	hasũ
01-NKK	Ø	taun	maur	toa	usa	ura	asu
02-PNR	Ø	taun	mauri	toa	usa	naura	asuasu
03-PTW	Ø	taon	mauri	toa	usa	ura	asuasu
04-WNV	Ø	taon	mauri	toa	u:sa	naura	a:su
05-TSM	Ø	taon	nauri	to:	usa	ura	a:su
06-WNR	Ø	taoni	maur	toa	uha	ura	aho
07-ELI	Ø	taoni	mauri	toa	uha	ura	
08-KRP	Ø	taon	mauri	toa	usa	ura	a:su
09-WSI	Ø	taun	tenmauri	toa	usa	ura	asu
10-LND	Ø	taun	tanmauri	toa	usa	ura	
11-TIA	Ø	taun	nauri	toa	usa	ura	a:su
12-KTN	Ø	tauni	mauri	toa	eusa	e ura	e-a:su
13-PNS	Ø	taun	mauri	toa	usa	naura	asu
14-WNS	Ø	taun	mauri	toa	usa	ura	n-a:su
15-PTL	Ø	taun	nauri	toa	usa	aura	asu

Table 14. Reflexes of PNCV *ʔ.

5.2 The PNCV affricate *z

The PNCV affricate *z, which reflects both POc *s and *j (affricate), is regularly reflected as the affricate tʃ or similar (e.g. [ts], cf. discussion in §5.1.3 on *d), see Table 15.

language	reflexes	FALL	FISH	STAR	CHEW	MOUTH	PARROT	GREEN
POc		*[s,j]u [a,u]			*jamu, *jam ^w -i-		*sipi, *sipi[r]i, *siri	*[keja]keja; *keja-ka
PNCV		*zovi	*mazi (Clark 2009b)	*mwazoe (*mwazoi)	*zamwa	*zigo	*zivi(ro) (TR) PNCV *siviri	*keza
01-NKK	tʃ			matʃo	tʃam		tʃiβ	malcetʃon
02-PNR	tʃ	tʃoβi	ma:tʃe	matʃoa	tʃamotʃam		tʃiβ	hetʃ

03-PTW	tʃ	tʃoβ		matʃoe	tʃama		tʃiβ	m ^w atʃina
04-WNV	tʃ	tʃoβi	ma:ts	m ^w atʃoe	tsamaja		tʃiβi	hetʃa
05-TSM	tʃ	tʃoβi		m ^w atʃoe	tʃomia		tʃiβkara	etʃa
06-WNR	tʃ	tʃoβi	ma:tʃi				natʃiβi	e:tʃa
07-ELI	tʃ	tʃo:vi	na:tʃi				tʃivkara	katʃana
08-KRP	tʃ	tʃovi	matʃi				paratʃiβi	
09-WSI	tʃ	tʃoβi	ma:tsi				paratʃiβi	etsa
10-LND	tʃ	tʃoβi	ma:tʃi		tʃa:ma	tʃino-	tʃitʃiβiru	
11-TIA	tʃ	tʃoβi	ma:ts			tsino-		
12-KTN	tʃ	tʃoβi	ma:tsi			tʃino-	tʃiβiro	
13-PNS	tʃ	tʃoβi	na:tʃi		tʃam ^w ae		tʃiβkara	etʃa
14-WNS	tʃ	tʃoβi	ema:tʃ e			etʃino-	paratʃiβi	
15-PTL	tʃ	tʃoβi	matsi	matʃoe			natʃiβi	hetʃa
16-PTS	tʃ	tʃo:βi	matʃi	m ^w atʃoi	tsam ^w ae		tʃiβi	he:tsa
17-TLM	tʃ	tʃoβi	natsi		tʃamaj		tʃiβi	

Table 15. Reflexes of PNCV *z.

5.3 The PNCV nasals

5.3.1 PNCV *m

As can be seen in Tables 16a, 16b and on Map 6, PNCV *m (< POC *m) is reflected as [m] regularly and throughout except in varieties that have undergone the linguolabial shift, i.e. *m > *n > n (cf. Clark 2009b, Lynch 2020b, also §5.1.1 on *b above). Such varieties in the sample are 07_Elia_Jo, 13_Panas_Jo and 17_Portolry_Tolomako. In all three, the linguolabial shift happened regularly before *a, as can be seen by the first four examples in Table 16a. The *ma sequence has a high type frequency with around 20 additional examples of it in the supplementary material.

The data for *m before front vowels is much scarcer. There are no examples with a PNCV *mi sequence. However, PNCV *qamuyu ‘you (pl.)’ appears to have lost its first *u early on, which resulted in *m before an i-like segment (probably more precisely [j]), but reflected in the three contemporary varieties as [i]), and the linguolabial shift appears to have taken place in this context in all three varieties. Regarding *me sequences, there are two examples: *mea ‘tongue’ and *mena ‘ripe’; in this environment there is evidence of the linguolabial shift in 17, but not in 07,13.

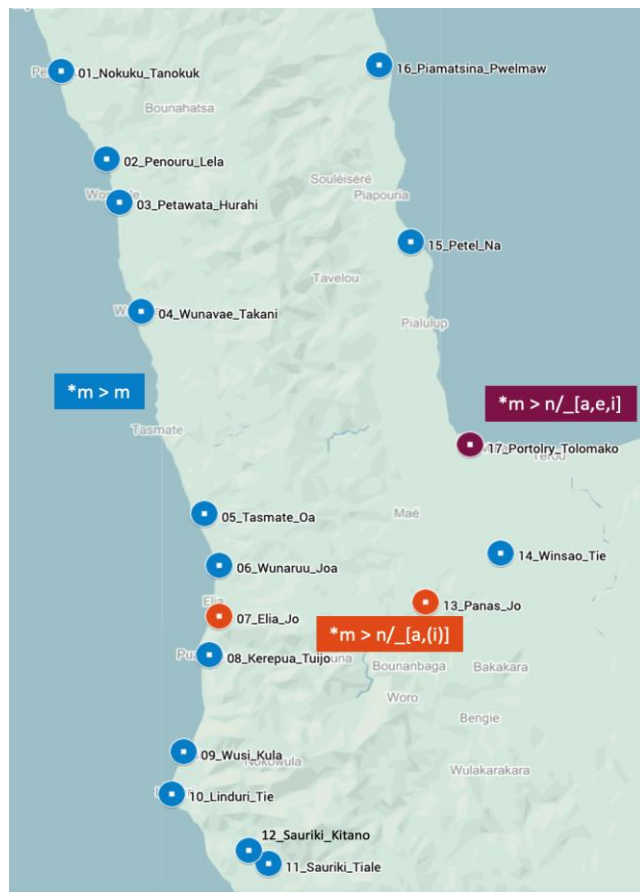
Finally, we see *m > n in a single example in 06_Wunaruu_Joa, *kanam* ‘we (1PL.EXCL)’ < *qama(m)i (highlighted in red in the table), which could be due to borrowing from a neighbouring variety. Alternatively, this could be a case of idiosyncratic *mVm > nVm dissimilation or -na- insertion, in the latter case the [m] in *kanam* reflecting the first *m of the proto-form.

language	reflexes	DIE	COCONUT	FIVE	WE_EXCL	EARTH QUAKE	MOSQUITO
POc		*mate	*matuqu	*lima			*ñamuk
PNCV		*mate	*matu-i	*lima	*qama(m)i	*muki	*namu-ki
01-NKK	m	mat		lima	omoto	mu	mok
02-PNR	m	mate	matui	lima	amato	mua	moh
03-PTW	m	ma:te	ma:tuj	lima		muj	mohı
04-WNV	m	mate	matui	lima	kamam	muj	mohı
05-TSM	m	mate	matui	lima	kamamu	muj	moʔi
06-WNR	m	mate	matui	lima	kanam	muj	moyı
07-ELI	n/ *[a,i], m	na:ti	natuj	lina	kanaj	muj	
08-KRP	m	mat	matuj	lima	kamaj	moj	
09-WSI	m	ma:ti	matui	lima	kamaj	mui	
10-LND	m	ma:te	matui	lima	kamaw	muj	muj
11-TIA	m	tenmate	matui	lima	kami	muj	muj
12-KTN	m	tanmate		lima	komao	muj	muj
13-PNS	n/ *[a,i], m	na:t	natui	lina	kanaj	moj	
14-WNS	m	ma:te	matui	lima	i ^u gamau	muj	
15-PTL	m	mate	matui	lima	ʔamam	namoj	mohi
16-PTS	m	ma:te	namatui	lima	kamamo	namuj	mohi
17-TLM	n/ *[a,e,i], m	nate	netuj	lina	kanam	muj	mox

Table 16a. Reflexes of PNCV *m (continues as Table 16b).

language	reflexes	ORANGE (CITRUS)	TONGUE	RIPE	CHIEF	YOU_PL
POc		*molis	*[kara-]maya-	*mena (Clark 2009a)		
PNCV		*moli	*mea	*mena	*moli (TR)	*qamuyu
01-NKK	m	mol	meme-	men	moli	om-to / emiu
02-PNR	m	molu	me-	mena	moli	am-to
03-PTW	m	mol pej	meme-	mena	mol	am-to
04-WNV	m	moli	meme-	tamena	moli	kamiu
05-TSM	m	moli	meme-	mena	moli	kamju
06-WNR	m	moli	meme-	mena	moli	kamiu
07-ELI	n/ *[a,i], m	moli	mimi-	mene	moli	kaniu
08-KRP	m	moli	mimi-	mene	moli	kamiu
09-WSI	m	moli	mimi-	mene	moli	kamiu
10-LND	m	moli	meme-	mena	mol	kamiu
11-TIA	m	moli	meme-		moli	kamiu
12-KTN	m	moli	meme-		moli	komiu
13-PNS	n/ *[a,i], m	moli	mimi-	mene	mol	kaniu
14-WNS	m	moli	meme-	mena	moli	^u gamiu
15-PTL	m	moli	meme-	mena	moli	ʔamiu
16-PTS	m	namol	meme-	mena	moli	kamiu
17-TLM	n/ *[a,e,i], m	mol	nene-	nena	moli	kaniu

Table 16b. Reflexes of PNCV *m (cont.).



Map 6. Reflexes of PNCV *m.

5.3.2 PNCV *m^w

The type frequency of PNCV *m^w (< POC *m^w) is low and the data is somewhat patchy. However, from the examples in Tables 17a, 17b we see some relatively clear patterns: *m^w is reflected consistently as plain [m] in 10,11,12,14,15,17 and consistently as [m^w] in 13,16. In the rest of the varieties, we see a mix of both, with [m^w] apparently more frequent than [m] in most of them. Given the relatively small amount of data and the tendency for labiovelarisation to be obscured during less careful speech, for it to be subject to sociolinguistic variation, and the fact that this may be an ongoing change (cf. §5.1.2, §5.8), a more precise explanation of this variation would require more data and analysis. Map 7 presents the observed patterns geographically.

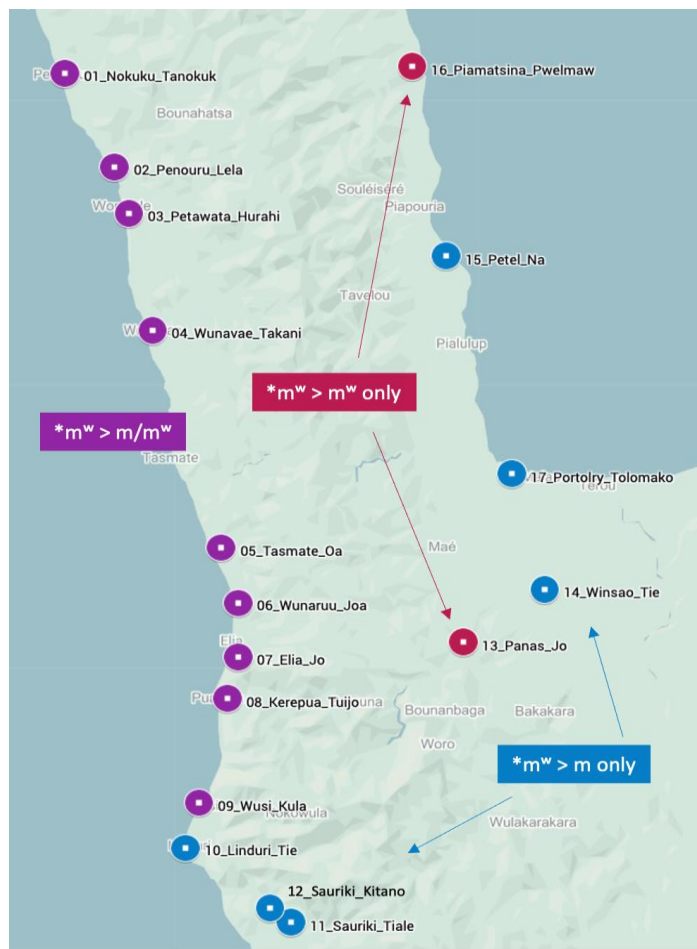
The [n] reflex in variety 17 *nape*- ‘liver’ appears to be due to an idiosyncratic *m^w > *m change that preceded the linguolabial shift (cf. §5.3.1).

language	reflexes	LIVER	CHILD	HOUSE	GARDEN	STAR	SUN
POC			*mera- mera	*Rumaq	*quma	PSV *-mwa(s,j)au (Clark 2009a)	
PNCV		*mwabwe	*mwara *mwera	*yumwa	*ʔumwa	*mwazoe (*mwazoi)	*mwaso (Clark 2009a)
01-NKK	m, m ^w	map ^w e-				matʃo	
02-PNR	m ^w , m	map ^w e				matʃoa	m ^w aso
03-PTW	m, m ^w	map ^w e-		ima	um ^w a	matʃoe	m ^w aso
04-WNV	m ^w , m	map ^w e-		im ^w a	wuma	m ^w atʃoe	m ^w aso
05-TSM	m ^w , m	map ^w e-		im ^w a		m ^w atʃoe	m ^w aso
06-WNR	m ^w , m	ma:pe			um ^w a		m ^w aso
07-ELI	m ^w , m	majm ^w ap ^w e	m ^w arusu		uma		m ^w aso
08-KRP	m ^w , m	m ^w ap ^w e-	maruhu		uma		m ^w aho
09-WSI	m ^w , m	m ^w ap ^w e-	maruhu		um ^w a		m ^w aho
10-LND	m	ma:pe-	natmera	ima	uma		
11-TIA	m	mape-	mera	ima	uma		
12-KTN	m	mape-	natimera	ima	uma		
13-PNS	m ^w	m ^w ap ^w e-	m ^w arusu		um ^w a		m ^w aso
14-WNS	m	mape-	netmera	eima	uma		
15-PTL	m	majmape			uma	matʃoe	maso
16-PTS	m ^w	m ^w ap ^w e-		im ^w a	um ^w a	m ^w atʃoi	m ^w aso
17-TLM	m	nape-			uma		maso

Table 17a. Reflexes of PNCV *m^w (continues as Table 17b).

language	reflexes	REEF	WORK	CHEW	SKY	YELLOW	WARM
POC		*m ^w aloq		*jamu, *jam ^w -i-			
PNCV		*mwalo	*ʔumwa	*zamwa	*mwazoe (*mwazoi)	*m ^w azina (TR)	m ^w am ^w am ^w a (TR)
01-NKK	m, m ^w	mal		tʃam		motʃin	
02-PNR	m ^w , m	m ^w alo		tʃamotʃam		m ^w atʃinā	
03-PTW	m, m ^w	m ^w alo		tʃama		m ^w atʃina	
04-WNV	m ^w , m	m ^w alo		tsamaja		matsina	
05-TSM	m ^w , m	m ^w alo		tʃomia	matʃaraʔa	m ^w atʃina	
06-WNR	m ^w , m	m ^w alo				m ^w atʃina	m ^w am ^w am ^w a
07-ELI	m ^w , m	m ^w alo	wa-um ^w a			matʃina	m ^w am ^w am ^w a otʃo
08-KRP	m ^w , m	m ^w alo				m ^w atʃina	m ^w am ^w am ^w a
09-WSI	m ^w , m	m ^w alo					m ^w am ^w am ^w a
10-LND	m		umauma	tʃa:ma		matʃina	
11-TIA	m					ta-matsina	
12-KTN	m					ta-matsina	ta-mamama
13-PNS	m ^w	m ^w alo	ʔa-um ^w a	tʃam ^w ae		m ^w atʃina	m ^w am ^w am ^w a otʃo
14-WNS	m					ta-matʃina	ta-mamama
15-PTL	m	malo				matʃina	
16-PTS	m ^w	m ^w alo		tsam ^w ae		m ^w atsina	
17-TLM	m		va-um	tʃamaj		matʃina	

Table 17b. Reflexes of PNCV *m^w (cont.).

Map 7. Reflexes of PNCV $*m^w$.

5.3.3 PNCV $*n$

PNCV $*n$ (< POC $*n$ and $*ñ$) remained unchanged in all positions in all varieties, as evidenced by the cognate sets in Table 18 and the supplementary material. I have highlighted an example from 09 where the $[n]$ reflex is followed by what I analyse as an epenthetic $[d]$ before $[r]$, resulting in a surface representation that is similar to the NDR sound, see §5.1.3 on $*d$ for a further discussion of this example.

language	reflexes	DAY	BIRD	YEAR	SMELL	EAT	LAKE	I
POC		$*rani$	$*manuk$	$*taun$	$*bo-na$	$*kani$		
PNCV		$*rani$	$*manu-ku$	$*taʔu-ni$	$*boni$ ($*buni$)	$*kani$	$*nobu$	$*nau$
01-NKK	n	rɒn	man	toun	ponpon	kɒnkɒn		nɒ
02-PNR	n	ran	mano	taun	poni	hanhan	nopũ	nao
03-PTW	n	ran	mano	taun	ponpon	han	nopu	nao
04-WNV	n	moran	manu	taun	ponina	ʔanʔan	nop	nao
05-TSM	n	ran	mano	taon	ponpon	ʔanʔan	nopu	naw
06-WNR	n	ran	mano	taon	ponponis	ʔaneyane		nao

07-ELI	n	rani	nano	taon	pono	ʔanʔan		nao
08-KRP	n	rani	mano	taoni	ponopono	ʔanʔan		nao
09-WSI	n	rani	ma:nu	taoni	pon ^d rone	ʔanʔani		nau
10-LND	n	ra:ni	manu	taon	ponaga	anani	no:pu	nao
11-TIA	n	rani	manu	taun	ponana	ko-anan	nopu	nao
12-KTN	n	taran	manu	taun	ponana	ko-anan	no:pu	nao
13-PNS	n	ran	na:no	taun	ponoina	ʔanʔan	nop	nao
14-WNS	n	tarani	ma:no	tauni	ponana	koanan	eni ^m pu	nao
15-PTL	n	ran	mano	taun	punina	han		nao
16-PTS	n	ran	mano	taun	ponpon	hani	nopu	nao
17-TLM	n	ran	nanu	taun	pon	vyan		nao

Table 18. Reflexes of PNCV *n.

5.3.4 PNCV *g

The PNCV velar nasal *g (< POc *ŋ) is reflected as [n] in most varieties (Tables 19a, 19b, Map 8), with the following exceptions and reservations:

- Fricative reflexes: In 15_Petel_Na, *g is regularly reflected as [h] and in 17_Portolry_Tolomako – as a velar fricative. Note that the latter variety might contrast velar and glottal fricatives (cf. §5.1.6 and Meyerhoff 2009: 5).
- Velar nasal reflexes: In 16_Piamatsina_Pwelmau the nasal velar has been retained regularly. The only counterexample is *senabulu* ‘ten’ (highlighted in Table 19b), which might be explained by contact with neighbouring varieties. In 03_Petawata_Hurahi, [n] and [ŋ] reflexes appear to have free distribution. Considering the reported historical ties between these communities (§2), it is possible that in 03 there may be a change in progress where [ŋ] is replaced by [n], likely due to contact with west coast varieties, and especially 01.
- Reflexes of word-initial *g: Most of the cognate sets in Tables 19a, 19b and the supplementary material involve PNCV *g in word-medial position. The only two sets with *g in word-initial position are WHEN (*gaisa) and TOOTH (*gado ‘gums’). In the former case, there is evidence of metathesis whereby initial *ga > *ag, which discards this as a useful example of word-initial *g. In the latter case, the words highlighted in grey in Table 19b appear indeed to be reflexes of *gado, as evidenced by the regular reflexes of *d (cf. §5.1.3). However, *g is not reflected by overt phonological material, which may mean either that in this particular case *g was dropped early on, or that *g is “regularly” reflected as Ø word-initially, at least in the relevant varieties. This remains an open question.

language	reflexes	HEAR	BAG	WIND	NIGHT	CRY
POC		*[l,r]oŋon; *roŋor-	*taŋa	*aŋin, cf. *laŋit 'sky, weather' (Clark 2009a)	*boŋi	*tanis-i-
PNCV		*rogo	*taga	*lagi	*bogi	*tagi-si (*dagi-si)
01-NKK	n	rono	tan	lɒn		
02-PNR	n	rono	tanõ	lan		
03-PTW	n, ŋ	roŋoe	tana	laŋ		taŋ
04-WNV	n	ronoja	tana	lan	poni	
05-TSM	n	ronoja	tana	lan	pon	
06-WNR	n	rone	tana	lan	pon	tan
07-ELI	n	rone	ta:na	lan	poni	
08-KRP	n	rone	tana	lan	poni	
09-WSI	n	rone	tana	la:ni	poni	
10-LND	n	pa:βrono	ta:na	laɟlan	po:ni	ta:ni
11-TIA	n	paβrono		lan	pon	-ta:ni
12-KTN	n	korono	ta:na	laelan	pon	-tani
13-PNS	n	rone	tana	lan	pon	
14-WNS	n	kurono	ta:na	elaɟlan	taponi	ta:ni
15-PTL	h	rohoj	taha	lahi	pohi	tahi
16-PTS	ŋ	roŋoe	taŋa	laŋi		taŋi
17-TLM	y/x	royoj	taxa	layi	poyi	tayi

Table 19a. Reflexes of PNCV *g (continues as Table 19b).

language	reflexes	EAR	TEN	WHEN	TOOTH
POC		*taliŋa	*sa ŋa puluq	*[i] ŋaican, *ican; *[qa]na-ŋ[a]ican	*ŋado-
PNCV		*daliga	*sagavulu	*gaisa	*gado
01-NKK	n	ɒnla-	sonoβul	p ^w anes	
02-PNR	n	anla-	sanaβul		
03-PTW	n, ŋ	anla-	senaβul	p ^w anesa	
04-WNV	n	kalna-		lanisa	
05-TSM	n	kalna-	sanaβul	anisa	
06-WNR	n	tʃalina-	sanaβulu	anisa	
07-ELI	n	tʃalina-	sanaβulu		
08-KRP	n	tʃalina-	hanaβul	ʔeniha	
09-WSI	n	tʃalina-	hanaβulu		
10-LND	n	kalna-	sanaβul	nanisa	ako-ga
11-TIA	n	kalna-	sanaβul	ananisa	ako-ka
12-KTN	n	kalna-	sanaβul	an:isa	a:ko-ka
13-PNS	n	tʃalina-	sanaβul		
14-WNS	n	kalna-	sanaβulu	anisa	e-ako-ka
15-PTL	h	tʃalha-		hisa	
16-PTS	ŋ	nd ralŋa-	senaβulu	ŋesa	
17-TLM	y/x	tʃaliya-		iyisa	

Table 19b. Reflexes of PNCV *g (cont.).



Map 8. Reflexes of PNCV *g.

5.4 The PNCV fricatives

5.4.1 PNCV *v

The PNCV labial fricative, denoted as *v (and descending from POc *p), likely had in that proto-language different labial fricative allophones (in mostly free distribution), as is the case with labial fricatives in many Vanuatu languages today. In the 17 varieties, *v is reflected mostly as a labial fricative or [w] (Tables 20a, 20b). When annotating the data, I have generally tried to transcribe labial fricatives narrowly in order to be able to detect any contrasts between them. This is methodologically challenging, as the auditory cues that differentiate between them are not always clear (Jongman 2024). I have also tried to rely on visual cues during data collection to distinguish between bilabial and labiodental fricatives. With these caveats in mind, the following generalisations can be made based on the data:

- By far the most common phonetic realisation in reflexes of *v is the bilabial voiced [β]. The other labial fricatives [v, f, φ] are also attested, although much less frequently. Their distribution appears to be mostly free. I have therefore used <β> as a cover symbol for the fricative reflexes.
- In all varieties, except 17_Portolry_Tolomako, there are also a few examples where *v surfaces as [w]. Such examples are significantly fewer than those with fricative reflexes. They also tend to appear in a handful of the around 40 cognate sets with *v, as exemplified in Table 20b. The last two columns in Table 20b represent the frequently colexified concepts FEATHER and HAIR, which here reflect the same PNCV etymon *vulu. There is a weak tendency for [w] to appear in HAIR and [β] in FEATHER (highlighted in yellow). This is insufficient evidence for contrast between β and w, especially given the colexification in the proto-languages and in most other contemporary varieties in the dataset. There is no evidence of contrast in the rest of the data either. There is a clear tendency for [w] to appear before rounded vowels, although there are exceptions so it is difficult to establish any patterns of complementary distribution between w and β.
- There are a handful of examples of a Ø reflex of *v, two of which are highlighted in red in Table 20b. These could be cases of w-loss as a result of a conditioned *tw>t change in these varieties, less careful speech and/or a general tendency for lenition.

language	reflexes	FOUR	PLANT (v)	RAT	LONG	PARROT	CRAB
POC		*vati			*b ^[w] arapu	*sipi, *sipi ^[R] i, *siri	*kape
PNCV		*pat; *pati	*lavo	*karivi	*baravu	*zivi(ro) (TR) < *siviri	*kave
01-NKK	β, (w)	βati	laβo	keriβ	peruβ	tʃiβ	ka:β
02-PNR	β, (w)	vati	laβoa	hariβ	paraβ	tʃiβ	ha:βe
03-PTW	β, (w)	vat	lawoe	hariβ	paraβ	tʃiβ	haβe
04-WNV	β, (w)	-βat	laβoja	ʔariβ	paraβ	tʃiβi	iyaβe
05-TSM	β, (w)	βati	laβoja	ʔariβi	paraβu	tʃiβkara	ʔaβe
06-WNR	β, (w)	βati	laβoʔ	ʔariβi	paraβu	natʃiβi	xaβe
07-ELI	β, (w)	va:ti	laβoʔe	ariβi	paraβu	tʃivkara	ʔaβe
08-KRP	β, (w)	βati	laβo	ʔariβ	paraβu	paratʃiβi	aβe
09-WSI	β, (w)	βa:ti	la:βo	ʔariβi	paraβu	paratʃiβi	aβe
10-LND	β, (w)	-βat	la:βo	arivi	paraβ	tʃitʃiβiru	a:ve
11-TIA	β, (w)	-va:t	-laβo	ariβi	taparaβ		aβe
12-KTN	β, (w)	-va:ti	-laβlaβo	ariβi	-paraβu	tʃiβiro	a:ve
13-PNS	β, (w)	βa:ti	laβo	ariβi	parawi	tʃiβkara	a:βe
14-WNS	β, (w)	-βa:ti	-laβo	ariβi	paraβo	paratʃiβi	e a:βe
15-PTL	β, (w)	φati	laβoe	hariβi	paraβu	natʃiβi	haβe
16-PTS	β, (w)	βa:ti	laβoe	hariβi	paraβu	tʃiβi	haβe
17-TLM	β	vati	laβoj	hariβi	taraβ	tʃiβi	yaβe

Table 20a. Reflexes of PNCV *v (continues as Table 20b).

language	reflexes	TURTLE	HILL	BUY	FEATHER	HAIR
POC				*poli	*pulu[-pulu]	*pulu[-pulu]
PNCV		*ʔavua	*vuti	*voli (*vuli)	*vulu	*vulu
01-NKK	β, (w)	oβu	wut	woli	wulu-	wul
02-PNR	β, (w)	awuh	wut	woli	wulu-	wulu-
03-PTW	β, (w)	aβua	wuti	wolia	βulũ	wul
04-WNV	β, (w)	aβua	naβuti	wulia	βulu-	wulu-
05-TSM	β, (w)	aβua	wuti	wolia	βulu-	βulu-
06-WNR	β, (w)	aβua	patiwuti	woli	wulu-	wulu
07-ELI	β, (w)	βua	p ^w atwuti	vol	wulu-	wulu-
08-KRP	β, (w)	βua	p ^w atut	woli	βulu-	wulu-
09-WSI	β, (w)	βua	p ^w atwuti	woli	wulu-	wulu-
10-LND	β, (w)	wua	patwut	wol=wol	fulu-	φulu-
11-TIA	β, (w)	wua	patwuti	-βoli	βulu-	wulu-
12-KTN	β, (w)	wua	patwut	-βoli	wulu-	wulu-
13-PNS	β, (w)	wua	p ^w atut	βoli	βulu-	βulu-
14-WNS	β, (w)	e βua	patwuti	-βoli	βulu-	βulu-
15-PTL	β, (w)	aβua	wuti	wolia	wulu-	wulu-
16-PTS	β, (w)	naβua	wuti	βolia	βulu-	wulu-
17-TLM	β	aβua	βut	βolia	βulu-	βulu-

Table 20. Reflexes of PNCV *v (cont.).

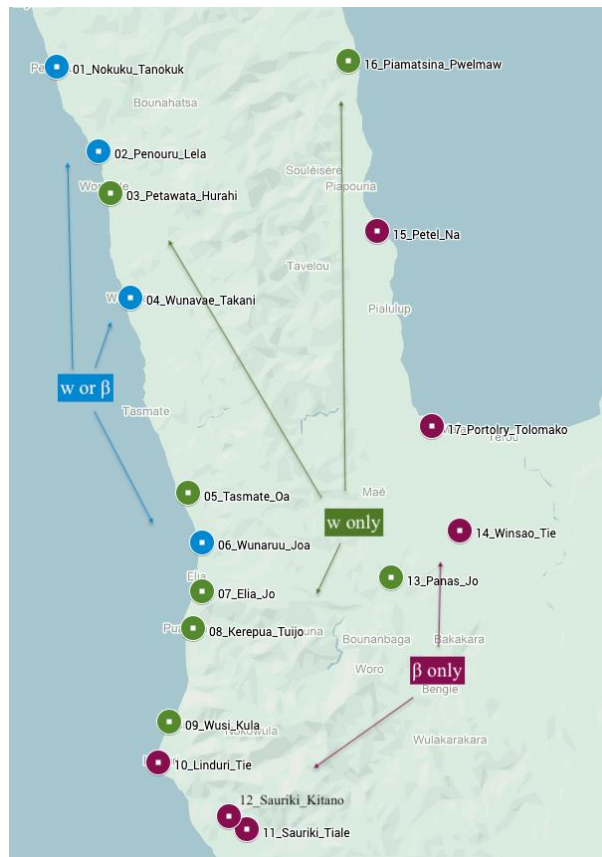
5.4.2 PNCV *v^w

Labiovelarisation as a secondary articulation has not been retained in any of the varieties in the reflexes of PNCV *v^w (< POC *p^w and *w). The data for these reflexes is scarce, see Table 21. In the southernmost varieties (10,11,12,14) and the two Big Bay varieties 15,17, the reflex of *v^w is consistently the plain bilabial fricative [β], while in varieties 03,05,07,08,09,13,16, it is consistently [w]. In the remaining four varieties (01,02,04,06), it is mostly [w], with some examples of [β]. See Map 9 for the geographic distribution of the patterns. It is difficult to make more precise statements considering the low type frequency of *v^w, but there is no evidence of complimentary distribution or contrast between β and w.

A note is due on the last two columns of Table 21. Here these are analysed as two different historical pathways from the same etymon. In the second one, *-na was added to the PNCV stem *vwa ‘fruit’, likely the 3SG possessive/construct suffix *-na (< POC *ña, Lynch et al. 2002: 76), which may have been reanalysed as a univerbated *v^wana, to which the -na suffix was added again in some of the varieties. The forms for varieties 07,09 could alternatively be analysed as wa-na, in which case they would be more accurately placed in the preceding column.

language	reflexes	“HERMAPHRODITE” PIG (NARAVE)	MOUTH	SAY	NINE	FRUIT	FRUIT
POc			*p ^{wa} ; *p ^{wa} p ^{wa} q	*wara, *warai-	*siwa	*puaq	
PNCV		*ravwe	*vwavwa	*vwara	*sivwa	*vwa	*v ^{wa} +*na (TR)
01-NKK	w,β	raβ	wɔwɔ?		tʃiwa		
02-PNR	w,β	raβe	wowa?		tʃiwa		
03-PTW	w	rawe	wawaw		tʃiwa		
04-WNV	w,β	naraβe	wawam				wana-na
05-TSM	w	narawe	wawana		tʃiwa		wana-na
06-WNR	w,β	naraβe	βaβaku	waraj			wana-na
07-ELI	w	rawe		werej			wana
08-KRP	w	rawe		wurej			wana-na
09-WSI	w	rawe	wawana	werej			wana
10-LND	β	ra:βe		varae		βae-na	
11-TIA	β	rave		-βarae			
12-KTN	β	raβe				βae-na	
13-PNS	w	ra:we		werej			wana-na
14-WNS	β	e ra:βe		daβaraj		βae-na	
15-PTL	β	tʃurevju	βaβam				βana-na
16-PTS	w	narawe	wawaku		tsiwa		
17-TLM	β	narave	vavam	βaraija			vona-na

Table 21. Reflexes of PNCV *vw.



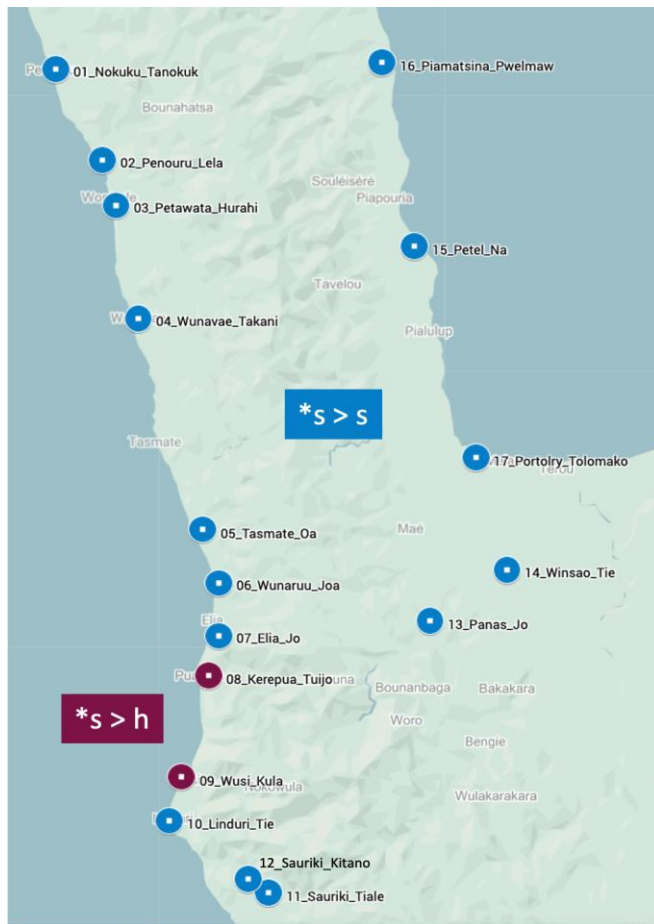
Map 9. Reflexes of PNCV *vw.

5.4.3 PNCV *s

PNCV *s (< POc *s, *j and *c) is reflected as [s] in all varieties except in 08_Kerepua_Tuijo and 09_Wusi_Kula, where it is regularly reflected as [h] (Table 22, Map 10). In the 30+ cognate sets, there are no exceptions to this rule. The speakers of 08,09 also report self-awareness of this difference and the h-reflexes in these varieties might act as shibboleths, i.e. markers of identity. In the form *heʔe-* ‘name’ < PNCV *kisa in variety 08, it may seem like *s is reflected as a glottal stop. However, the glottal stop is a regular reflex of *k in this language, which means that *heʔe-* is likely the product of metathesis.

language	reflexes	BREAST	KINGFISHER	SEA	RAIN	SWIM	NAME
POc		*susu[-]	*[s,j]iko	*tasik	*qusan	*lo-loso[p] *losop-i-	*[q]aca-; *i[c,s]a-
PNCV		*susu	*siko	*tasi	*ʔusa	*loso-vi	*kisa > *asa
01-NKK	s	sus	sik	tɒs	us	lolos	kise-
02-PNR	s	susu-	siho	tas	usa	lolos	his
03-PTW	s	susu	siho	tas	usa	loloso	hise
04-WNV	s	susu	siyo	ta:si	usa	loloso	ise-
05-TSM	s	susu-	siʔo	tasi	usa	loloso	ise
06-WNR	s	susu	nasifo	ta:si	u:sa		yise
07-ELI	s	susu	siʔo	tas	usa	loloso	ese-
08-KRP	h	huhu-	hiʔo	tahe	uha	lololoho	heʔe-
09-WSI	h	huhu-	hiʔo	tahi	uha	loloho	ehe
10-LND	s	susu-	sio	ta:si	usa	loloso	e:se-
11-TIA	s	susu-	sio	ta:si	usa	ko-loloso	esa-
12-KTN	s	susu-	sio	ta:si	usa	loloso	ese-
13-PNS	s	susu-	siʔo	ta:s	usa	lolos	ese-
14-WNS	s	susu-	e sio	ta:si	eusa	-loloso	ese-
15-PTL	s	susu-	nasiho	tasi	usa	loloso	hise-
16-PTS	s	sysy-	siho	ta:si	usa	loloso	hisej-
17-TLM	s	susu-	siyo	tas	usa	lolos	hise-

Table 22. Reflexes of PNCV *s.



Map 10. Reflexes of PNCV *s.

5.5 The PNCV lateral *l

PNCV *l (< POC *l) is reflected as /l/ in all varieties without exception in the 40+ cognate sets involving this proto-sound, see Table 23 and the supplementary material.

language	reflexes	PLANT (v)	FIVE	SWIM	THREE	STONE
POC			*lima	*lo-losop	*tolu	
PNCV		*lavo	*lima-	*loso-vi	*tolu	*sule (Clark 2009b–Proto Santo)
01-NKK	l	laβo	lima	lolos	tolu	sul
02-PNR	l	laβoa	lima	lolos	tolu	sule
03-PTW	l	lawoe	lima	loloso	tolū	sule
04-WNV	l	laβoja	-lima	loloso	-tolu	sule
05-TSM	l	laβoja	lima	loloso	tolu	sule
06-WNR	l	laβo?	lima		tolu	sule
07-ELI	l	laβoʔe	lina	loloso	tolu	sula
08-KRP	l	laβo	lima	lololoho	tolu	hula
09-WSI	l	la:βo	lima	loloho	tolu	hula
10-LND	l	la:βo	-lima	loloso	-tol	su:le

11-TIA	l	-laβo	-lima	-loloso	-tol	sule
12-KTN	l	-laβlaβo	-lima	loloso	-tolu	esule
13-PNS	l	laβo	lina	lolos	tolu	sula
14-WNS	l	-laβo	-lima	-loloso	ta-tolu	sule
15-PTL	l	laβoe	lima	loloso	tolu	sule
16-PTS	l	laβoe	lima	loloso	tolu	sule
17-TLM	l	laβoj	lina	lolos	tolu	sule

Table 23. Reflexes of PNCV *l.

5.6 The PNCV rhotics

5.6.1 PNCV *r

PNCV *r (< POc *r and *R) is reflected as [r] in all varieties without exception in the 30+ cognate sets involving this proto-sound, see Table 24 and the supplementary material.

language	reflexes	DAY	TWO	SMELL	CRAYFISH	RAT
POc		*rani	*rua	*[l,r]oŋon; *roŋoR-	*quraŋ	
PNCV		*rani	*rua	*rogo	*ʔura	*karivi
01-NKK	r	rɒn	rua	rono	ur	keriβ
02-PNR	r	ran	rua	rono	ura	hariβ
03-PTW	r	ran	rua	roŋoe	ura	hariβ
04-WNV	r	moran	-rua	ronoja	naura	ʔariβ
05-TSM	r	ran	rua	ronoja	ura	ʔariβi
06-WNR	r	ran	rua	rone	naura	ʔariβi
07-ELI	r	rani	rua	rone	ura	ariβi
08-KRP	r	rani	rua	rone	ura	ʔariβ
09-WSI	r	rani	rua	rone	ura	ʔariβi
10-LND	r	ra:ni	-rua	pa:βrono	ura	arivi
11-TIA	r	rani	-rua	paβrono	ura	ariβi
12-KTN	r	ta-ran	-rua	-rono	ura	ariβi
13-PNS	r	ran	rua	rone	ura	ariβi
14-WNS	r	ta-rani	-rua	-rono	e ura	ariβi
15-PTL	r	ran	rua	rohoj	naura	hariβi
16-PTS	r	ran	rua	roŋoe	ura	hariβi
17-TLM	r	ran	rua	royoj	aura	hariβi

Table 24. Reflexes of PNCV *r

5.6.2 PNCV *R

The data involving the other PNCV rhotic, *R (< POc *R), are relatively scarce; all attested examples appear in Table 25. *R appears to have been regularly lost in all 17 varieties. In OCTOPUS, it could potentially be analysed as being reflected by [w,u,o], but these are much more likely reflexes of *u, the preceding *R having (as is regular) been lost. Regarding COCONUT CRAB, Clark (2009a: 99) comments that “[a] majority of languages regularly lose both medial consonants [in PNCV *daweRu], leaving an unusual sequence of three vowels, which makes it difficult to tell

which developments are regular”; in the present data, *w appears to be regularly reflected by [u,o,w] or Ø, and *R is most likely lost. In most varieties, there appears to be reflexes of an additional final syllable, which I reconstruct as *ki or *ke, with regular reflexes of *k (cf. §5.1.6).

In the last two columns there are counterexamples (highlighted) where *R appears to be reflected by [r]. It is difficult to make any conclusions based on these few examples. They could be due to borrowing, irregular reflexes or other developments. In any case, this is not enough evidence to posit any hypotheses about a conditional development. A similar lack of regularity for the reflexes of *R has been observed also in other part of Vanuatu and other branches of the Oceanic family (Geraghty 1990, François 2011).

language	reflexes	OCTOPUS	BONE	BITE	HIGH TIDE	COCONUT CRAB	STING RAY	FOG
POC		*kurita	*suri[-]	*karat-i	*[u]ruap		*pari	
PNCV		*kuRita	*suRi	*kaRa-ti	*Rua	*daweRu (Clark 2009a) / *daweRu-ki (TR)	*vaRi	*ma-Ravu
01-NKK	Ø	kwit	sui-	kɔtmoti	tɔs u	ow		
02-PNR	Ø	hwit	sui-	hatmoti	tas ua	ae		
03-PTW	Ø	hwita	sui-		ua	a:ihi		
04-WNV	Ø	wita	sui-	atusia	tas mo-ua	kawi		
05-TSM	Ø	wita	sui-	ʔatmotia	tasi ua	kauʔi		
06-WNR	Ø	nawita	sui	xati tusia	tasi oua	tʃaihi		
07-ELI	Ø	uita	sui	et tus	ta:si oua	ʔaβe tʃaiʔi		
08-KRP	Ø	wita	hwi-		tah ua	kaui		
09-WSI	Ø	uita	huj-		tah eua	kaui		
10-LND	Ø	wita	sui-	a:ti	tas ta-ua	kaoi		
11-TIA	Ø	wita	sui-	-atrot	tas ta-ua	kaoi		mara β
12-KTN	Ø	wita	sui-	-atro:to	tas ta-ua	kaoi		
13-PNS	Ø	wita	sui-		tas oua	kawi		
14-WNS	Ø	nawita	esui-	-atroto	tas ta-samua	kaoe		
15-PTL	Ø	nahwita	sui-	hat tusia		tʃaihi		
16-PTS	Ø	hoita	sui-	hatmotia		nd raehi	vari	
17-TLM	Ø	ita	sui	hatia		tʃaixi	vari	

Table 25. Reflexes of PNCV *R.

5.7 The PNCV approximants

5.7.1 PNCV *w

All cognate sets involving PNCV *w (< POC *w) are in Table 26. Two general patterns stand out:

- *w > Ø word-initially as evidenced by the first four cognate sets, with the last cognate set (PENIS) being a counterexample.

- In word-medial position, *w is reflected as [w] or otherwise its [+rounded] feature is retained as [o,u]. I use capital <W> as a cover symbol for these in the reflexes (sound) column.

A discussion of a few exceptions or caveats is in order here:

- *w appears to have been lost in a few varieties in cognate set COCONUT CRAB (highlighted in yellow), where W is expected (see also the discussion of this concept in §5.6 above).
- In 01_Nokuku_Tanokuk, the W reflex appears to be absent where expected in the words highlighted in green. However, this variety is known to have very recently lost final vowels (which may have been /o/ or /u/). In fact, such a dropped final [u] is evidenced in *βakamerv* ‘left’, where former *a is known to have been rounded before a high vowel (§4.1, Clark 2009a: 27).
- In a few cases (highlighted in pink), it appears that *w may be reflected as [e] or [j]. These may be irregular reflexes or may be due to some additional phonological material that is not evidenced in the proto-forms.
- In 03_Petawata_Hurahi, *halo* ‘eight’ most likely contains an epenthetic [h], cf. §5.1.7.
- My reconstruction *wasi ‘rope’ reflects POC *waroc. POC *c > PNCV *s is regular (Clark 2009a: 16).
- PNCV *mawiri ‘left’ appears to have undergone early metathesis.

language	reflexes	PALOLO WORM	EIGHT	CANOE	ROPE	SKY	COCONUT CRAB
POC		*weli	*walu	*waga	?*waroc	*masawa(n,ŋ)	
PNCV		*weli	*walu	*waqa	*wasi (TR)	*masawa (*masava)	*daweRu (Clark 2009a) / *daweRu-ki (TR)
01-NKK	Ø, W	el	ɲlu		ɲs	mase	ow
02-PNR	Ø, W	eli	alu		as	maseã	ae
03-PTW	Ø, W	e:li	halo		a:s		a:ihi
04-WNV	Ø, W				ja:s		kawi
05-TSM	Ø, W	eli	alu		asi		kau?i
06-WNR	Ø, W				asi		tʃaihi
07-ELI	Ø, W	eli			a:si		ʔaβe tʃai?i
08-KRP	Ø, W	eli			ahi		kaui
09-WSI	Ø, W				ahɪ		kaui
10-LND	Ø, W	eli		aka	a:si		kaoi
11-TIA	Ø, W	eli			a:s		kaoi
12-KTN	Ø, W	tʃaroj eli			eas		kaoi
13-PNS	Ø, W	eli			na a:s		kawi
14-WNS	Ø, W				ease		kaoe
15-PTL	Ø, W				asi		tʃaihi
16-PTS	Ø, W		-alu		a:si	masau	nd raehi
17-TLM	Ø, W				a:s		tʃaixi

Table 26a. Reflexes of PNCV *w (continues as Table 26b).

language	reflexes	LEFT	SHARK	YAWN	ROOT	KAVA	PENIS
POC		*mawiri, *mauri-	*bakewa	*[ma]mawap	*kawari	*kawa (PROc)	PCP *waqe 'leg' (Clark 2009a)
PNCV		*mawiri	*bakewa	*mawa-va	*kawa- ri		*wae
01-NKK	Ø, W	βakamerɔ	pace	kɔn mɛmaw		ka	
02-PNR	Ø, W	vaha marao	paheo	hanmao	huari-	hohoe	
03-PTW	Ø, W	maraw	paheo	hanmauɸ	hwari-	hae	wae
04-WNV	Ø, W	maraw	payeo	?an mamaw	wari-	ae	
05-TSM	Ø, W	maraw	pa?eo	?an mamauβi			wae-na
06-WNR	Ø, W		paheo	xan mamaoβi	hoe-		
07-ELI	Ø, W	naraw	pa?oe	?an nanaoβi			wae-na
08-KRP	Ø, W	marau-	pa?oe	?an mamaoβ			
09-WSI	Ø, W	marao	pa?oe	?an mamauβi			
10-LND	Ø, W	marau-	paeo				
11-TIA	Ø, W	marau-	paoe				
12-KTN	Ø, W	marau-	paoe				
13-PNS	Ø, W	naraw	pa?oj	?an nanauβ			
14-WNS	Ø, W	marau-					
15-PTL	Ø, W	maraw		han mamauvi	hoe-	haj	
16-PTS	Ø, W	marau		han mamauβi	huari-	hae	wai-m
17-TLM	Ø, W	narao		yananauv	xoe-		

Table 26b. Reflexes of PNCV *w (cont.).

5.7.2 PNCV *y

There are only a handful of cognate sets featuring the PNCV palatal approximant *y (< POc *y) in the data set and all of them appear in Table 27. The scarcity of the data means that conclusions can be made with less confidence. However, the following patterns emerge:

- *y can be either lost or otherwise be reflected as a front vowel [i,e] or palatal approximant [j].⁹
- In EEL, in most varieties, *y appears to have merged with the last *a to result in [e], except for in varieties 07,13 where the entire last syllable appears to have been lost.
- In YOU (PL), the first *u appears to have been lost, and *y retained as [i]. In varieties 01,02,03, the forms *om-to* and *am-to* reflect truncated complex forms that involve reflexes of PNCV *tolu 'three'.

⁹ [j] may in some cases be considered a phonetic surface representation of an underlying front vowel, especially in unstressed positions.

language	reflexes	SUN	WOOD	HOUSE	YOU_PL	EEL
POC			*kayu	*Rumaq		*malayo *maraya (PROc)
PNCV		*yal	*kayu	*yumwa	*qamuyu	*maraya
01-NKK	Ø, (i)	al	kɔ		om-to/emiu	mera
02-PNR	Ø, (e)		haw		am-to	namarae
03-PTW	Ø, i/e		haw	ima	am-to	marae
04-WNV	Ø, i/e		haw	im ^w a	kamiu	marae
05-TSM	Ø, i/e		ʔaw	im ^w a	kamiu	namarae
06-WNR	Ø, i/e		ʔaw		kamiu	namarae
07-ELI	Ø, i		ʔaw		kaniu	nara
08-KRP	Ø, i/e		ʔaw		kamiu	marae
09-WSI	Ø, i/j		ʔaw		kamiu	maraj
10-LND	Ø, i/e	alo	wuj-ao	ima	kamiu	marae
11-TIA	Ø, i/e	alo	wi-ao	ima	kamiu	marae
12-KTN	Ø, i/e	mata-i alo	wi-ao	ima	komiu	marae
13-PNS	Ø, i		ʔaw		kaniu	nara
14-WNS	Ø, i/e	e-alo	wi-ao	e-ima	ᵝgamiu	e marae
15-PTL	Ø, i/e		haw		ʔamiu	marae
16-PTS	Ø, i/e		haw	im ^w a	kamiu	marae
17-TLM	Ø, i/e		ayao		kaniu	narae

Table 27. Reflexes of PNCV *y.

5.8 Discussion of the evolution of the consonant inventories

Firstly, Table 28 presents likely consonant phonemic inventories for the varieties based on the data.

PNCV	p	p ^w	t	k	ʔ	m	m ^w	n	ŋ	β	s	x ⁷	h	ʈ	r	nd _r	l	w	j ⁸
01_Nokuku_Tanokuk																			
02_Penouru_Lela																			
03_Petawata_Hurahi									4										
04_Wunavae_Takani					2									6					
05_Tasmate_Oa					2									6					
06_Wunaru_Joa																			
07_Elia_Jo						3													
08_Kerepua_Tuijo											5								
09_Wusi_Kula											5								
10_Linduri_Tie																			
11_Sauriki_Tiale																			
12_Sauriki_Kitano																			
13_Panas_Jo						3													
14_Winsao_Tie																			
15_Petel_Na																			
16_Piamatsina_Pwelma																			
17_Portolry_Tolomako	1					3													

Table 28. Summary of the likely consonant inventories of the contemporary varieties. Black boxes mean likely phonemes; grey boxes mean likely marginal phonemes or lower confidence; the following notes provide detail:

- 1) Before /o,u/
- 2) Mostly in initial position in 04, and medial position in 06
- 3) Only before vowels other than /a/ in 07,13 and only before /o,u/ in 17
- 4) Not a regular reflex of *g, likely an ongoing process of merger with /n/

- 5) Only in borrowings, including the exonym *Wusi*.
- 6) Mostly in medial position in 04, in initial position in 06
- 7) *x~h* distinction uncertain in other varieties, more likely in 17 based on previous literature (cf. §5.1.6)
- 8) Based on the available data it is difficult to argue whether a surface [j] is a phoneme on its own or a representation of an underlying front vowel

A summary of the reflexes of all PNCV consonants can be found in Table 29. A discussion follows.

PNCV	*b	*b ^w	*d	*t	*q	*k	*ʔ	*v	*v ^w	*s	*z	*m	*m ^w	*n	*g	*l	*r	*R	*w	*y
01_Nokuku_Tanokuk	p	p ^w /(p)	ʔ/∅	t	ʔ/∅	k	∅	β, (w)	w, β	s	tʃ	m	m, m ^w	n	n	l	r	∅	∅, W	∅, (i)
02_Penouru_Lela	p	p ^w /(p)	ʔ/∅	t	ʔ/∅	h/(h)	∅	β, (w)	w, β	s	tʃ	m	m ^w , m	n	n	l	r	∅	∅, W	∅, (e)
03_Petawata_Hurahi	p	p ^w /(p)	ʔ/∅	t	ʔ/∅ (k)	h	∅	β, (w)	w	s	tʃ	m	m, m ^w	n	n, ɲ	l	r	∅	∅, W	∅, i/e
04_Wunavae_Takani	p	p ^w /(p)	k	t	k	ʔ/∅, (h/ʔ)	∅	β, (w)	w, β	s	tʃ	m	m ^w , m	n	n	l	r	∅	∅, W	∅, i/e
05_Tasmate_Oa	p	p ^w	k/(ʔ)	t	k	ʔ/(∅)	∅	β, (w)	w	s	tʃ	m	m ^w , m	n	n	l	r	∅	∅, W	∅, i/e
06_Wunaru_Joa	p	p/p ^w	tʃ	t	k	ɣ/x/h/, ʔ/∅	∅	β, (w)	w, β	s	tʃ	m	m ^w , m	n	n	l	r	∅	∅, W	∅, i/e
07_Elia_Jo	p	p ^w	tʃ	t	k	ʔ/∅	∅	β, (w)	w	s	tʃ	n, m	m ^w , m	n	n	l	r	∅	∅, W	∅, i
08_Kerepua_Tuijo	p	p ^w /(p)	tʃ	t	k	ʔ/∅	∅	β, (w)	w	h	tʃ	m	m ^w , m	n	n	l	r	∅	∅, W	∅, i/e
09_Wusi_Kula	p	p ^w /(p)	tʃ	t	k	ʔ/∅	∅	β, (w)	w	h	tʃ	m	m ^w , m	n	n	l	r	∅	∅, W	∅, i/j
10_Linduri_Tie	p	p	k	t	k	∅	∅	β, (w)	β	s	tʃ	m	m	n	n	l	r	∅	∅, W	∅, i/e
11_Sauriki_Tiale	p	p	k	t	k	∅	∅	β, (w)	β	s	tʃ	m	m	n	n	l	r	∅	∅, W	∅, i/e
12_Sauriki_Kitano	p	p	k	t	k	∅	∅	β, (w)	β	s	tʃ	m	m	n	n	l	r	∅	∅, W	∅, i/e
13_Panas_Jo	p	p ^w	tʃ	t	k	ʔ/∅	∅	β, (w)	w	s	tʃ	n, m	m ^w	n	n	l	r	∅	∅, W	∅, i
14_Winsao_Tie	p/m ^w	p	k	t	k/ʔk/ʔg	∅	∅	β, (w)	β	s	tʃ	m	m	n	n	l	r	∅	∅, W	∅, i/e
15_Petel_Na	p	p	tʃ	t	ʔ	h	∅	β, (w)	β	s	tʃ	m	m	n	h	l	r	∅	∅, W	∅, i/e
16_Piamatsina_Pwelmau	p	p ^w	nd _r	t	k	h	∅	β, (w)	w	s	tʃ	m	m ^w	n	ɲ	l	r	∅	∅, W	∅, i/e
17_Portolry_Tolomako	t, p	p	tʃ	t	k	x/ɣ/h	∅	β	β	s	tʃ	n, m	m	n	ɣ/x	l	r	∅	∅, W	∅, i/e

Table 23. Summary of the reflexes of all PNCV consonants in all varieties. Within each column where substantial variation in reflexes is observed, differential highlighting is used for better visualisation of shared innovations and retentions.

Arguably the most dramatic change involved the reorganisation of the plosive inventory. PNCV's two rows of plosives based on a prenasalisation distinction has collapsed in all varieties into a single row of plosives (with no prenasalisation distinction¹⁰). This was likely supported by the lack of ***p* and ***pw*, which made the inventory unbalanced and allowed for greater variation within the articulation of labial plosives. The labiovelarisation contrast was lost in some varieties and has been maintained (to different degrees) in others. In 17, as a result of the linguolabial shift, **p* and **t* merged before **/a, e, i/*.

The loss of the prenasalisation distinction did not simply involve a merger of **d* with **t* and of **q* with **k*. Varieties 06,07,08,09,13,15,17 have retained the place of articulation of **d* but changed the manner of articulation to an affricate. This affricativisation may have taken place before or after loss of prenasalisation, i.e. **d[n̥d] > *n̥dʒ/*n̥tʃ > tʃ*, or **d[n̥d] > *d (> *t/*dʒ) > tʃ*. In the rest of the varieties (except 16) the reflex of **d* is *k* or *ʔ/∅*, which means that a change in place of articulation was involved. All these varieties have merged **d* and **q* (cf. Clark 2009a: 26). In the varieties where **k* changed to a fricative, glottal stop or zero, and **q* (and **d* where they merged) changed to *k* (likely though intermediate stages, possibly involving a **g[g]* stage), the two changes were likely involved in a push/pull chain. An interesting case is variety 01, which

¹⁰ However, note the prenasalised allophones in 14 (§5.1.1, §5.1.4), and the retention of a prenasalised NDR-sound in 16 (§5.1.3), which likely patterns with trills, not plosives, at least in the synchronic phonology.

corresponds to Clark's (*ibid.*) Nokuku variety, where *k was retained but *d and *q merged to a glottal stop/zero. This suggests that the change from *d and *q to glottal stop/zero did not involve a stage during which they merged with *k.¹¹ Instead, it likely involved a stage where a plain voiced *g contrasted with *k, i.e. there appears to have been a voicing distinction somewhere along the plosive evolution of this variety, i.e. *d[n̥d] > *ŋg/*d > *g[g] > ʔ/Ø and *q[ŋg] > *g[g] > ʔ/Ø. This scenario would mean a relatively early *d~*q merger. Considering the historical connections between variety 01 and varieties 02,03 and even beyond, it is then unlikely that the *g[g] > ʔ/Ø change involved an intermediate *k. Instead, a fricative stage can be hypothesised, likely a (voiced) velar fricative, which would have contrasted with the glottal fricative that *k may have already changed into in the relevant varieties. Table 24 summarises the proposed ordered changes. Alternatively, in varieties 02,03 there could have been an intermediate stage involving a *k on the pathway to ʔ/Ø, assuming that *k had already changed to *h by the time. However, such a development involving relatively early stages that did not affect variety 01, is a less parsimonious explanation.

	*d, *q	*k
1.	*d[n̥d] > *ŋg/*d > *g[g] *q[ŋg] > *g[g] (merger of *d and *q)	*k > k (01) / h (02,03)
2.	*g[g] > *ɣ > ʔ/Ø	

Table 24. A summary of the proposed hypothesis for the development of *d, *q, *k in varieties 01,02,03.

Another systemic change involves labiovelarisation. It was lost in varieties 10,11,12,14,15,17. This is true for all three PNCV labiovelarised consonants (*b^w, *v^w, *m^w) and resulted in the plain variants of these sounds. Varieties 10,11,12 of the west coast are geographically contiguous and they have historical connections to variety 14 (§2). Varieties 15,17 are also spoken close to each other on the Big Bay coast, so this could also be an areal effect, either independent of the west coast area or as part of a larger area. This remains an open question pending further data collection and analysis.

In the rest of the varieties, labiovelarisation as a secondary articulation was retained in reflexes of *b^w and *m^w to at least some extent. In varieties 13,16, it is retained without exception where expected. Additionally, in 05,07, *b^w is consistently reflected by a labiovelarised [p^w]. Otherwise, we see some (free) variation between labiovelarised and plain reflexes. Labiovelarisation can have low phonetic salience, especially in rapid or less careful articulation, which means that the variation here may be due to data scarcity or idiolectal effects. Furthermore, it is possible that in some varieties there may be an ongoing loss of labiovelarisation, or that

¹¹ Assuming it is unlikely that PNCV *k underwent changes that were later reversed.

labiovelarisation has sociophonetic constraints, as attested in other languages of Vanuatu such as Unua (Pearce 2015: 22) and Ahamb (Rangelov 2020: 56), cf. also Clark's (2009a: 11) observation discussed in §5.1.2.

The situation is similar with $*v^w$, where labiovelarisation is preserved (as primary articulation) in the same varieties that preserved labiovelarisation in plosives and nasals, again, to varying degrees.

Turning to the linguolabial shift, which was discussed in detail in §5.1.1 and §5.3.1, it affected more substantially variety 17 ($*b > t/_*[a,e,i]$, $*m > n/_*[a,e,i]$) and less so varieties 07,13 ($*m > n/_*[a,i]$). In the latter two varieties this development has high salience in that speakers reported being very aware of it and claim that this is a prominent distinction of their varieties from other neighbouring varieties.

Another phonological shibboleth is the $*s > h$ shift in varieties 08,09.

Finally, we saw that PNCV $*g[\eta]$ merged with $/n/$ in most varieties except for 03,16 (spanning the mountain chain) where a velar nasal reflex is observed and in the Big Bay coastal varieties 15,17 where it became a fricative.

Summarising the above observations, the following patterns in how the varieties cluster together can be identified based on shared innovations.

- Varieties 10,11,12,14 share five innovations ($*d > k$, $*q > k$, $*k > \emptyset$, $*g > n$, and loss of labiovelarisation). Varieties 04,05 also share the first four innovations with them.¹² However, only $*d > k$ is an innovation exclusive to these six varieties.
- Varieties 07,08,09,13 share four innovations ($*d > \text{tʃ}$, $*q > k$, $*k > \text{ʔ}/\emptyset$, $*g > n$), none of them being exclusive to them. Furthermore, this cluster is divided into two pairs based on whether they underwent the linguolabial shift (07,13) or the $*s > h$ shift (08,09).
- The Big Bay coastal varieties 15,17 share the following innovations: $*d > k$, $*k > \text{fricative}$, $*g > \text{fricative}$, and loss of labiovelarisation. Only one of these innovations is exclusive to this pair ($*g > \text{fricative}$). Furthermore, only variety 17 underwent the linguolabial shift.
- Varieties 01,02,03 share the innovations $*d > \text{ʔ}/\emptyset$ and $*q > \text{ʔ}/\emptyset$. This merger is exclusive to this cluster of varieties. However, note that $*q > \text{ʔ}$ is also observed in 16.

Finally, I compare the results in Table 23 with those of Tryon (1976) for the relevant varieties (cf. §2). Tryon (1976: 11-50) lists correspondences to POc phonemes based on Grace (1969).¹³ Overall, the results largely overlap, with the following notable exceptions: (1) Tryon does not report labiovelarised p^w as a reflex of the POc labiovelarised plosive (Grace's $\langle \eta p \rangle$) in any of the varieties, while m^w is reported only in Piamatsina; (2) Tryon's main reflex of POc $*k$ ($>$ PNCV $*k$) for Nokuku is $\text{ʔ}/\emptyset$, while in the current analysis we see strong support for k ; (3) Tryon did not

¹² One caveat is that they also allow $[\text{ʔ}]$ as a reflex of $*k$. We could speculate that the $\text{ʔ}/\emptyset$ variation in these two varieties reflects an ongoing process of loss of the glottal stop, which has already been completed in 10,11,12,14.

¹³ The two analyses are based on different reconstructions. The current analysis is based on both PNCV (Clark 2009a) and POc (Ross et al. 1998, 2007, 2008, 2011, 2016, 2023) reconstructions, while Tryon relies only on Grace's older POc reconstructions. However, a general comparison is both possible and informative.

identify an open back rounded vowel in Nokuku, but instead listed [o] and [u] as possible reflexes of *a, the latter appearing “[b]efore or near u” (Tryon 1976: 36).

6. Summary and conclusion

This paper discussed the phonological history of 17 varieties of Santo based on new data. For many of the varieties, these are the first data known to have been collected. The discussion was organised with the PNCV phoneme inventory as a starting point. One of the aims of this paper was to shed light on processes of phonological evolution of languages in this understudied part of Santo.

The vowel inventories did not undergo any significant changes, besides the addition of a sixth back open vowel in variety 01. The consonant inventories underwent much more substantial changes, including a considerable reorganisation of the plosive inventories of all varieties. Other innovations include loss of labiovelarisation, the velar nasal merging with /n/ in most varieties, an idiosyncratic *s > h shift in two varieties, and the well-attested linguolabial shift, which was an intermediate stage resulting in a partial merger of labial and coronal plosives and nasals in three varieties.

The overlapping shared innovations do not provide a clear picture for subgrouping. The current Glottolog classification of the languages Tie and Tiale (varieties 10,11) as belonging to the same primary branch of West Santo is largely confirmed, with varieties 12,14 also seemingly belonging to this group. More tenuous ties were observed between this group and varieties 04 and 05, the latter corresponding to the language Tasmate, which is within a different subgroup under existing classifications. Tasmate also does not appear to share many innovations with its sister varieties 01 (Nokuku) and 16 (Piamatsina) as per the current Glottolog classification. Kula (varieties 08,09) clusters with the previously undocumented varieties 07,13. The previously undocumented variety 15 appears to cluster somewhat with Tolomako (17). The further investigation of the classification and language/dialect status of these varieties is in preparation by the author. Ultimately, such investigations would benefit immensely from new data from other areas of this understudied island, which is a natural laboratory of linguistic diversity.

Supplementary material

The supplementary material can be found at the following link:
<https://doi.org/10.5281/zenodo.22762180>

The following files are included:

- An ODS file (open in LibreOffice, Microsoft Excel or similar) called Rangelov-2026--17 SantoVarieties-PhonHist-DATA.ods containing the following spreadsheets:
 - ‘complete data set’ – contains all lexical data and metadata
 - ‘data set for phonological historical analysis’ – contains all lexical data whose cognate coding has been checked

- 25 sheets with collections of cognate sets demonstrating the reflexes of each of the 25 PNCV phonemes
- A PDF file [Rangelov-2026--17 SantoVarieties-PhonHist-SupplMat-MethodsEtc.pdf](#) with a description of the data and methods related to data collection, annotation, organisation and cognate coding.

Acknowledgements

The data used in this paper was collected in 2022, 2023 and 2025. Data collection was funded by the Department of Linguistic and Cultural Evolution at the Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany. Fieldwork and research were carried out under an ethics approval by the Max Planck Society and a research agreement with the Vanuatu Cultural Centre. I would like to thank the Vanuatu Cultural Centre, the Tabwemasana Council of Chiefs, the Santo Sunset Environment Network, the Area Administrators of West Coast and Northwest Santo areas, and all members of the communities that I visited, and especially Allan Taman and family, Sabrina Taman, Chief Ben Frank Palla and family, Chief Noel Sari, Deacon James Linrow, George Frank, Elder Vaiaken Aulu, Tani Marklin, Joshua Lewalev, Stephen Tavue, Chief Richard Poi, Charly Lapiho, Chief Taitas Rojo, Stefan Pipite, Chief Song Vanua, Pune Ropwel, Peka James, and Chief Puha Lapan. I also want to thank Russell Gray, Mary Walworth, Miriam Meyerhoff, Ross Clark, Elizabeth Pascal, Ben King, Isaac Stead, Russel Barlow, Jakob Halfmann, Natalie Korobzow, and Hans Schmidt for their encouragement, company, and/or comments on an earlier draft of this paper. I am grateful to the editor and reviewer for their constructive comments and suggestions, which helped improve the manuscript. All remaining errors are my own.

References

- Barbour, Julie, Gayleen Tarosa & Kanaueha Wessels. 2022. The language varieties of Malua Bay and Espiegle Bay (Tepërav), Malekula: A comparative study of consonants. *Language and Linguistics in Melanesia* 40. 60–107.
- Barbour, Julie. 2024. The phonological history of consonants in Neverver, a Western Malekula language. *Language and Linguistics in Melanesia* 42. 204–236.
- Clark, Ross. 2006. *Notes on Big Bay Language (Tolomako)*. Unpublished manuscript.
- Clark, Ross. 2009a. **Leo Tuai: A comparative lexical study of North and Central Vanuatu languages*. Canberra: Pacific Linguistics.
- Clark, Ross. 2009b. *Espiritu Santo as a “linguistic area.”* Paper presented at the 11th International Conference on Austronesian Linguistics (11ICAL), Aussois, France.
- Crowley, Terry. 2000. The language situation in Vanuatu. *Current Issues in Language Planning* 1(1). 47–132. <https://doi.org/10.1080/14664200008668005>

- Evans, Nicholas. 2022. Pushing the boundaries: Marginal phonemes and dialogic interaction. *Russian Journal of Linguistics* 26(4). 995–1011. <https://doi.org/10.22363/2687-0088-32349>
- François, Alexandre. 2011. Where *R they all? The geography and history of *R-loss in Southern Oceanic languages. *Oceanic Linguistics* 50. 140–197. <https://doi.org/10.1353/ol.2011.0009>.
- François, Alexandre, Michael Franjieh, Sébastien Lacrampe & Stefan Schnell. 2015. The exceptional linguistic diversity of Vanuatu. In Alexandre François, Sébastien Lacrampe, Michael Franjieh & Stefan Schnell (eds.), *The languages of Vanuatu: Unity and diversity*, 1–21. Canberra: Asia-Pacific Linguistics.
- Geraghty, Paul. 1990. Proto Eastern Oceanic *R and its reflexes. In Jeremy H.C.S. Davidson (ed.), *Pacific Island languages: Essays in honour of G.B.Milner*, 51–94. School of Oriental and African Studies, University of London/ University of Hawaii Press.
- Grace, George W. 1969. A Proto-Oceanic finder list. *Working Papers in Linguistics* 1. 38–84. Department of Linguistics, University of Hawai'i.
- Hammarström, Harald, Robert Forkel, Martin Haspelmath & Sebastian Bank. 2026. *Glottolog* 5.3. Leipzig: Max Planck Institute for Evolutionary Anthropology. <https://doi.org/10.5281/zenodo.18840935> (Available online at <http://glottolog.org>, Accessed on 2026-03-25.)
- Jongman, Allard. 2024. Phonetics of Fricatives. In Mark Aronoff (ed.), *Oxford Research Encyclopedia of Linguistics*. Oxford University Press. <https://doi.org/10.1093/acrefore/9780199384655.013.1086>
- Ladefoged, Peter & Ian Maddieson. 1996. *The sounds of the world's languages*. Oxford, UK: Blackwell.
- Lindstrom, Lamont. 1994. Traditional cultural policy in Melanesia (Kastom polisi long kastom). In Lamont Lindstrom & G. M. White (eds.), *Culture-kastom-tradition: Developing cultural policy in Melanesia*, 67–81. Institute of Pacific Studies, University of the South Pacific.
- Lynch, John. 2016. The phonological history of Iaa. *Language and Linguistics in Melanesia* 33. 53–77.
- Lynch, John. 2019a. The phonological history of Naman, a Western Malakula language. *Language and Linguistics in Melanesia* 37. 21–41.
- Lynch, John. 2019b. The phonological history of Nese, a Northern Malakula Language. *Language and Linguistics in Melanesia* 37. 61–82.
- Lynch, John. 2020a. The phonological history of Uripiv, an Eastern Malakula Language. *Language and Linguistics in Melanesia* 38.
- Lynch, John. 2020b. The bilabial-to-linguolabial shift in Southern Oceanic: A subgrouping diagnostic? *Oceanic Linguistics* 58(2). 292–323.
- Lynch, John & Terry Crowley. 2001. *Languages of Vanuatu: A new survey and bibliography*. (Pacific Linguistics, 517.) Canberra: Australian National University.
- Lynch, John, Malcolm Ross & Terry Crowley. 2002. *The Oceanic languages*. Richmond: Curzon.

- Meyerhoff, Miriam. 2009. *Aspects of the language spoken at Matantas, Big Bay (Vanuatu)*. Unpublished manuscript.
- Monnier, Paul. 1994. *Les lettres du Père Alphonse Ardouin (1880 – 1940)*. Port-Vila: Maison Mariste.
- Pearce, Elizabeth. 2015. *A grammar of Unua*. Berlin: De Gruyter Mouton.
- Rangelov, Tihomir. 2020. *A grammar of the Ahamb language (Vanuatu)*. PhD thesis. University of Waikato, New Zealand. <https://hdl.handle.net/10289/14038>
- Rangelov, Tihomir. 2023. The acoustic and articulatory properties of the prenasalised coronal trill in two Oceanic languages. In Radek Skarnitzl & Jan Volín (eds.), *Proceedings of the 20th International Congress of Phonetic Sciences, Prague 2023*, 3417–3421.
- Rangelov, T. 2024. *The NDR sound in Austronesian languages*. Paper presented at 16th International Conference on Austronesian Linguistics. Manila, Philippines. Manila. 2024-06-20 - 2024-06-24.
- Rangelov, Tihomir. 2026a. *Language density hotspots*. Zenodo. <https://doi.org/10.5281/zenodo.18647227>
- Rangelov, Tihomir. 2026b. Identifying, documenting, and describing marginal phonemes: A fieldworker's guide. *Acta Linguistica Academica* 73(2). 220-245. <https://doi.org/10.1556/2062.2025.01015>
- Rangelov, Tihomir, Eleanor Ridge & Lana Takau. 2025. Linguistics in Vanuatu 45 years after Independence. *Te Reo* 68(2). 31–78.
- Ridge, Eleanor. 2025. What does *lanwis* mean? Parallels between Indigenous languages of Vanuatu for young people in Mele Maat. *Te Reo* 68(2). 133–161.
- Ross, Malcolm, Andrew Pawley & Meredith Osmond. 1998. *The lexicon of Proto Oceanic: The culture and environment of ancestral Oceanic society. Volume 1: Material culture*. Canberra: Australian National University.
- Ross, Malcolm, Andrew Pawley & Meredith Osmond. 2007. *The lexicon of Proto Oceanic: The culture and environment of ancestral Oceanic society. Volume 2: The physical environment*. Canberra: Australian National University.
- Ross, Malcolm, Andrew Pawley & Meredith Osmond. 2008. *The lexicon of Proto Oceanic: The culture and environment of ancestral Oceanic society. Volume 3: Plants*. Canberra: Australian National University.
- Ross, Malcolm, Andrew Pawley & Meredith Osmond. 2011. *The lexicon of Proto Oceanic: The culture and environment of ancestral Oceanic society. Volume 4: Animals*. Canberra: Australian National University.
- Ross, Malcolm, Andrew Pawley & Meredith Osmond. 2016. *The lexicon of Proto Oceanic: The culture and environment of ancestral Oceanic society. Volume 5: People: body and mind*. Canberra: Australian National University.
- Ross, Malcolm, Andrew Pawley & Meredith Osmond. 2023. *The lexicon of Proto Oceanic: The culture and environment of ancestral Oceanic society. Volume 6: People: society*. Canberra: Australian National University.

- =====
- Taman, Allan, Sabrina Taman & Tihomir Rangelov. 2025. Language ecosystems in Western Santo. *Te Reo* 68(2). 186–245. <https://doi.org/10.5281/zenodo.18723278>
- Touati, Benjamin. 2014. *Description du sakao, langue du nord-est Santo (Vanuatu)*. Paris: Université de Paris IV-Sorbonne PhD thesis.
- Tryon, Darrell. 1976. *New Hebrides Languages: An internal classification*. (Pacific Linguistics: Series C, 50.) Canberra: Australian National University.
- Tryon, Darrell. 2010. The languages of Espiritu Santo, Vanuatu. In John Bowden, Nikolaus P. Himmelmann & Malcolm Ross (eds.), *A journey through Austronesian and Papuan linguistic and cultural space: Papers in honour of Andrew K. Pawley*, 283-290. Canberra: Australian National University.