

Discovering history through language: papers in honour of Malcolm Ross

Edited by Bethwyn Evans



Pacific Linguistics
Research School of Pacific and Asian Studies
The Australian National University

Published by Pacific Linguistics
Research School of Pacific and Asian Studies
The Australian National University
Canberra ACT 0200
Australia

Copyright in this edition is vested with Pacific Linguistics

First published 2009

National Library of Australia Cataloguing-in-Publication entry

Title: Discovering history through language : papers in honour
of Malcolm Ross / editor Bethwyn Evans.

ISBN: 9780858836051 (pbk.)

Notes: Bibliography.

Subjects: Ross, Malcolm.
Historical linguistics.
Anthropological linguistics.
Linguistic change.

Other Authors/
Contributors: Evans, Bethwyn.

Dewey Number: 417.7

Copyedited by Bethwyn Evans/Jeanette Coombes
Typeset by Jeanette Coombes
Cover design by Julie Manley
Photo of Lapita pottery fragment courtesy of Christophe Sand, New Caledonia
Printed and bound by Addcolour Digital Pty Ltd, Fyshwick, Canberra

6 *The languages of Vanikoro: three lexicons and one grammar*

ALEXANDRE FRANÇOIS

Cette île, toute petite qu'elle est, présente le singulier phénomène de plusieurs idiômes differens. (Gaimard 1833:338)

1 The paradox of Vanikoro languages¹

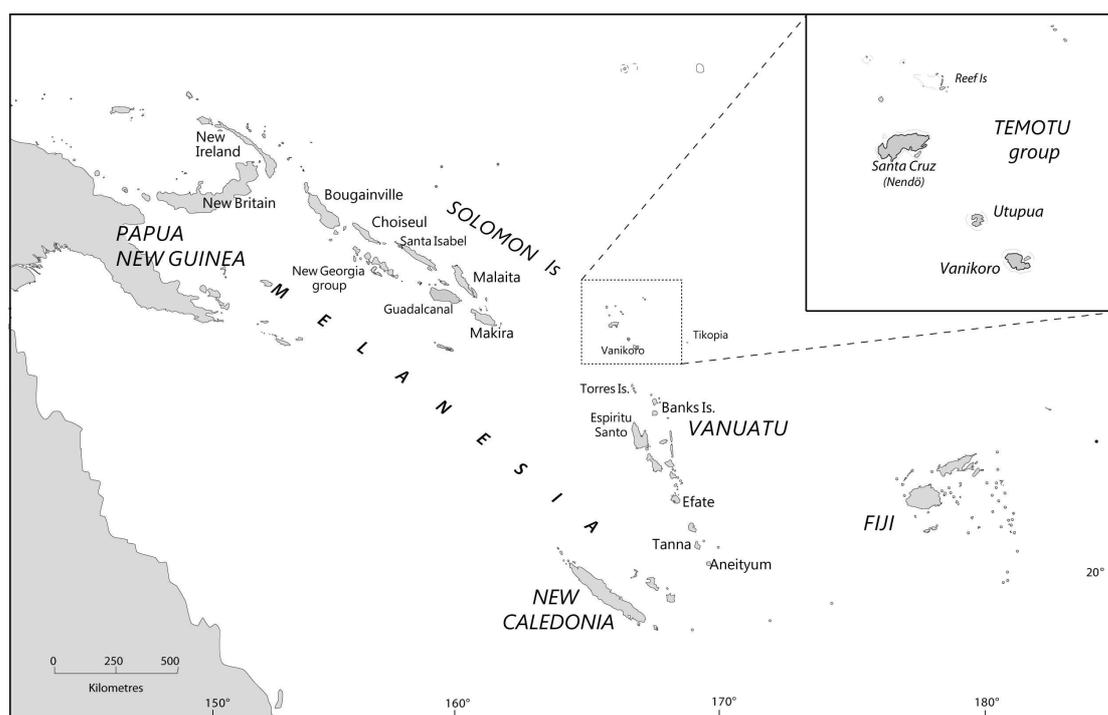
1.1 The languages of Vanikoro

With its 193 sq. km, Vanikoro is the second largest island in the small archipelago formerly known as the Santa Cruz Islands, and now often referred to as ‘Temotu’, after the official name of the easternmost province of the Solomon Islands (Map 1).

The province of Temotu is home to a variety of languages (Tryon 1994): three Polynesian, and nine non-Polynesian. The latter include three languages on Vanikoro, and three on Utupua — a total of six Oceanic languages which have long been understood to form a branch of their own (Tryon and Hackman 1983). The three remaining languages, known as the ‘Reefs-Santa Cruz’, were long deemed to be Papuan (Wurm 1976), but have recently been shown to be Austronesian (Ross and Næss 2007). More specifically, Ross and Næss have proposed to group all the non-Polynesian languages of the region into a single first-order subgroup of Oceanic, labelled ‘Temotu’. The latter would then split into two branches: Reefs-Santa Cruz (RSC) on the one hand, and Utupua-Vanikoro (UV)² on the other hand.

¹ My gratitude goes to Malcolm for having fostered my linguistic research on Melanesian languages, both through his writings and through our discussions, ever since we first met in 1997. His interest in my data has been a strong incentive for me to spend the years 2009 and 2010 at The Australian National University. My initial work on Vanikoro was facilitated by the Institut de Recherche pour le Développement of Nouméa and Association Salomon; by the French Centre National de la Recherche Scientifique; by Piet Lincoln; by Association ‘Banie’ and the traditional chiefs of the island; and by my Vanikoro helpers and friends — especially Stanley Repuamu, Ezekiel Prians, John Nabu. I am grateful to Bethwyn Evans and Andrew Pawley for their comments on earlier versions of this paper.

² Except for its new location in the POc tree, this UV branch coincides with the subgroup identified earlier as ‘Eastern Outer Islands’ (Tryon 1994, 1995).



Map 1: Location of Vanikoro in the Pacific

If Ross and Næss' hypotheses are correct, then Temotu constitutes a new branch of the Oceanic family tree whose history needs to be written. The present paper hopes to play its part in this endeavour, by presenting primary data and some discussion on the three indigenous languages of Vanikoro: Teanu (TEA), Lovono (LVN) and Tanema (TNM).

Published information on Vanikoro languages is still limited, but certainly not absent. In 1788, the island of Vanikoro was reached by the two frigates of the French navigator Jean-François de La Pérouse, and saw his fatal shipwreck — an event which was only understood a few decades later by the Irish navigator Peter Dillon (Dillon 1830). Another French officer, Dumont d'Urville, immediately organised an expedition in Dillon's wake. Among the abundant documentation produced by this second voyage (Dumont d'Urville 1830–1834), the French naturalist Gaimard (1833, 1834) compiled wordlists in the three languages of Vanikoro, a precious document on the linguistic situation of that time.³ Much more recently, other word lists have been compiled by Tryon and Hackman (1983), based on an extended version of the Swadesh basic vocabulary list. Darrell Tryon also wrote short grammatical accounts of Lovono (Tryon 1994:630–634), and of Teanu (Tryon 2002). Additionally, a short collection of Teanu texts was published in Tua and Lincoln (1979).⁴

Almost two centuries after Dumont d'Urville's expedition, the French Ministère de la Marine, together with Association Salomon and Institut de Recherches pour le Développement, organised another expedition called *Vanikoro 2005*, to find out about the fate of La Pérouse's ships and sailors. I was given the opportunity to play my part there as

³ I am much indebted to Piet Lincoln for allowing me easy access to Gaimard's documents.

⁴ The two languages Teanu and Lovono have been given varying names over time. Teanu was called *Tanéanou* by Gaimard, and *Buma* by Tryon. Lovono was called *Vanikoro* by Gaimard, *Vanikolo* by Ivens (1918), *Vano* by Tryon. See §1.2 below for a discussion of my naming proposals.

a linguist, documenting place-names and oral traditions, with a special interest in the islanders' stories that still remember so vividly the 1788 wreckage (François 2008a). On this occasion, I was also able to follow in Gaimard's footsteps, and record what I could of the three languages. One thing I realised was the urgency of this task, with both Lovono and Tanema remembered by only a handful of speakers.

1.2 A note on the history and geography of Vanikoro

The population of Vanikoro can be described at two different levels of observation.

A contemporary look would probably suggest just a binary divide between two communities, one Melanesian and one Polynesian. The latter is a group of about 300 settlers originating from Tikopia, a small island located about 200 km eastwards. Although they have been colonising the southern shores of Vanikoro for more than three centuries (see Dillon 1830), they tend to interact very little with the native population — except for the occasional land dispute. As their social network is still anchored in their Polynesian homeland, they remain predominantly monolingual in Tikopian, the Polynesian Outlier language spoken on Tikopia (Firth 1985). Apart from a few loanwords here and there, this recent colonisation does not show any major linguistic consequence, and will not be discussed further.

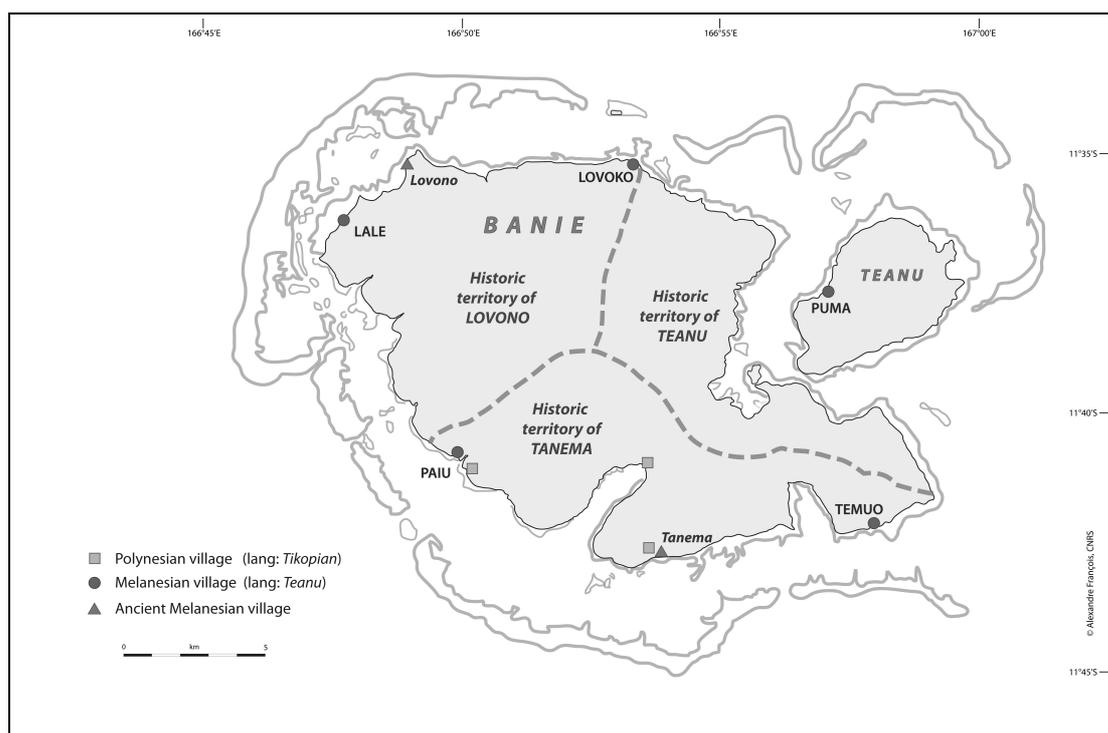
With about 600 individuals descending from the earlier inhabitants of Vanikoro, the Melanesians live today in six coastal villages scattered around the island: Puma, Temuo, Paiu, Lale, Lovono, Lovoko. People from these villages communicate through coastal canoeing, and now form a homogeneous society. This modern unity reportedly results from the action of the Anglican missionaries who christianised the region at the end of the 19th century. However, the first historical documents on Vanikoro (Dillon 1830; Dumont d'Urville 1830–1834), as well as the oral tradition of the islanders themselves, tell a different story: that of an island which used to be sharply divided into three distinct 'tribes' or chiefdoms. Each of these tribes was attached to a specific territory which they defended fiercely from one another, and which is still clearly delimited in people's memories, see Map 2.

Each tribe's name recalls a significant place of its own area:

- ***Teanu*** from the name of the northeast island of the Vanikoro group, where the village of *Puma* is also located⁵
- ***Lovono*** from the name of a village, also formerly known as *Vono* or *Vano*,⁶ northwest of the main island Banie
- ***Tanema*** an ancient village, also known as *Tetawo*, on the southern coast (in what is now *de facto* Polynesian territory)

⁵ The village name *Puma*, sometimes with the incorrect spelling *Buma*, has been used as another name for the language of Teanu (see footnote 4).

⁶ The Lovono name of this village of Vano/Lovono was apparently *Vana* or *Alavana*. In principle, one might want to use this autonym as a reference name for this language, yet this would be slightly artificial: the only term which is used today, even by the last speakers themselves, is the form *Lovono* in Teanu, the only surviving language.



Map 2: The populations of Vanikoro

Each of these three tribes had its own language, which survived up until the 20th century. Gaimard's (1833, 1834) wordlists cite 'Tanéma' and 'Vanikoro' (Lovono) on an equal footing with 'Tanéanou' (Teanu). Ivens (1918:155), in his translations of the prayer 'Our Father', illustrates Vanikoro with a language that can be identified as Lovono — which suggests it was still in a healthy state at the end of the 19th century. But since the pacification of the island, intermarriage amongst the three earlier communities has increased, and they have merged into a single society. Probably due to its earlier demographic lead, Teanu was adopted as the whole island's daily language, very much at the expense of the two other vernaculars. In 2005, Lovono and Tanema were still remembered by only a handful of elder speakers — five for Lovono, four for Tanema. Because these last speakers are now dispersed across the various Teanu-speaking villages of Vanikoro, they do not form anything that would resemble a speech community. Clearly, the shift to Teanu is now complete: Lovono and Tanema are now two moribund languages, with only a few years left to live.

1.3 Three closely related languages

1.3.3 Three aberrant Oceanic languages

For the linguist accustomed to other Oceanic-speaking areas, the three languages of Vanikoro can be disconcerting. This section will touch briefly upon the question of their genetic affiliation, before we examine the issue of their mutual relationship.

One conspicuous characteristic of Teanu, Lovono and Tanema is the degree to which they differ, both lexically and morphologically, from most other known Oceanic languages (François 2006), whether from the Solomons or from nearby Vanuatu. For example, these

three languages show relatively little retention of the lexicon reconstructed for Proto Oceanic. Many etyma, which tend to be otherwise widely preserved among Oceanic languages (Pawley 2007), have disappeared altogether from the lexicon of Vanikoro languages; e.g. *tama- ‘father’, *mate ‘die’, *kani ‘eat’, *kayu ‘tree’, *ikan ‘fish’, *sapa ‘what’, *jalan ‘road’, *susu ‘milk’, *pano ‘go’, *sake ‘up’, *sipo ‘down’. The morphology also shows a great deal of erosion. Thus the languages of Vanikoro show no trace whatsoever of the POc possessive affixes (*-gu, *-mu, *-ña...; *-qi...), of object pronouns, of the article *na, of common verbal affixes like the transitive suffix *-i, the applicative *-aki[n], the causative *pa[ka]-, and so on.

In sum, Vanikoro languages can be considered ‘aberrant’ (see Grace 1990; Pawley 2006) in comparison with most other Oceanic languages. This does not mean, however, that their Oceanic ancestry cannot be detected. Indeed, despite the high level of lexical replacement which evidently took place in their history, it is still possible to uncover some words whose similarity with Oceanic reconstructions is beyond doubt, and which could hardly be attributed to chance or borrowing — if only because their phonological correspondence patterns seem to follow some degree of regularity. Table 1 provides a small sample list of such words.⁷

Table 1: Some obvious Oceanic reflexes across the three languages

English	Teanu	Lovono	Tanema	POc etymon
‘bird’	<i>menuko</i>	<i>menuka</i>	<i>manuke</i>	*manuk
‘eye’	<i>mata</i>	<i>mala</i>	<i>ka\mae</i>	*mata
‘soul, spirit’	<i>ata</i>	<i>ala</i>	<i>ae</i>	*qata
‘ears’	<i>taña</i>	<i>mabe\neje</i>	<i>añe</i>	*taliña
‘house’	<i>moe ~ mwoe</i>	<i>moe</i>	<i>nalama</i>	*Rumaq
‘cold’	<i>medigo</i>	<i>meniña</i>	<i>mediña</i>	*ma ⁽ⁿ⁾ ri ⁽ⁿ⁾ riŋ
‘long, tall’	<i>biouro</i>	<i>beure</i>	<i>va\beura</i>	*barapu
‘Canarium nut’	<i>vo\ŋoro</i>	<i>ve\ŋere</i>	<i>vi\ŋara</i>	*[ka]ŋari
‘lie down’	<i>wene</i>	<i>enu</i>	<i>eno</i>	*qenop

1.3.2 Proto Vanikoro, the common ancestor

Not only can Teanu, Lovono and Tanema be individually linked to Proto Oceanic, but they can also be shown to form a set of three closely related languages. It is possible to suggest instances of exclusively shared innovations, thereby pointing to the likely existence of a common ancestor Proto Vanikoro. The following paragraphs therefore answer the question raised by Ross and Næss (2007:473), according to whom ‘no innovations define Vanikoro’.

⁷ The three Vanikoro languages have the same phoneme inventory, with five short vowels (i, e, a, o, u) and 19 consonants. The spelling conventions here adopted include *v*=/β/; *ñ*=/ɲ/; *j*=/ʝ/. Labiovelar consonants use digraphs with a *w*, and all prenasalised voiced stops are spelled without their nasal element, e.g. /^mb^w/ is *bw*, /^ŋg/ is *g*, etc.

In the phonological domain, one can cite the addition of a paragogic vowel, probably a schwa, after most word-final consonants, resulting in the retention of that consonant followed by a non-etymological vowel:

- (1) POC *manuk ‘bird’ > *manukə > TEA *menuko* ~ LVN *menuka* ~ TNM *manuke*;
 POC *maⁿriⁿriŋ ‘cold’ > *maⁿriŋə > TEA *medigo* ~ LVN *meniŋe* ~ TNM *mediŋa*;
 POC *wair ‘water’ > *wairə > TEA *ero* ~ LVN *wire* ~ TNM *nʌira*.

Another example of a sound change which is only found in Vanikoro languages is the occasional velarisation of *t to /k/ before a high back vowel /u/:

- (2) POC *patu ‘stone’ > *vatu > *vaku > *vakə > TEA *voko* ~ LVN/TNM *vaka*;
 POC *kutu ‘louse’ > *utu > *uku > *ukə > TEA *uko* ~ LVN/TNM *-uka*.

In the morphological domain, we will see (§3.1.2) that the three languages share the same structural collapse between certain non-singular personal prefixes.

Finally, many lexical items seem to be shared exclusively by these three languages. Table 2 proposes a set of possible lexical innovations in Proto Vanikoro — with tentative reconstructions of the most likely etymological forms.⁸

Table 2: Some putative lexical innovations in Proto Vanikoro

English	Teanu	Lovono	Tanema	Proto Vanikoro
‘moon’	<i>metele</i>	<i>mele</i>	<i>maloula</i>	*mataul(ə)
‘fish’	<i>namuko</i>	<i>namweka</i>	<i>namaka</i>	*nam ^(w) əkə
‘wood, tree’	<i>vilo</i>	<i>kuile</i>	<i>veila</i>	*v ^(w) eilə
‘taro’	<i>jeʋbute</i>	<i>bule</i>	<i>bue</i>	*bute
‘child’	<i>menu</i>	<i>melika</i>	<i>anuka</i>	*menuk(ə)
‘woman’	<i>emele</i>	<i>neme</i>	<i>me</i>	*nemel(ə)
‘name’	<i>eŋa</i>	<i>neŋe</i>	<i>niŋe</i>	*neŋə
‘who’	<i>ŋele</i>	<i>eŋe</i>	<i>ŋela</i>	*eŋel(ə)
‘inhabited land’	<i>kulumoe</i>	<i>kulamoe</i>	<i>kulama</i>	*kuləmaʔ(ə)
‘rubbish’	<i>ajekele</i>	<i>togale</i>	<i>soge</i>	*jogel(ə)
‘tell s.o.’	<i>viñi</i>	<i>puñi</i>	<i>puña</i>	*puñə
‘heavy’	<i>motoro</i>	<i>melure</i>	<i>mwaura</i>	*matur(ə)
‘dirty’	<i>sukiro</i>	<i>sukure</i>	<i>vatukura</i>	*sukir(ə)
‘stand’	<i>vio</i>	<i>pia</i>	<i>veo</i>	*piə
‘impede, protect’	<i>botoŋo</i>	<i>beloŋa</i>	<i>baŋa</i>	*batŋə
‘do again’	<i>tabo</i>	<i>lebu</i>	<i>abo</i>	*tabo

⁸ Due to the lack of solid regularity, in particular, in the domain of vowels, it is difficult to securely reconstruct any protoform based on the synchronic data. This is an important difference between the languages of Vanikoro and other parts of Oceania, where protoforms can be reconstructed based on modern reflexes. See Ross (1988) for the languages of western Melanesia, Lynch (2001) for southern Vanuatu, François (2005) for northern Vanuatu.

In order to ascertain these reconstructions and expand the list of Proto Vanikoro innovations, more needs to be known of the languages on the neighbouring island Utupua. But the evidence given here should be sufficient to establish that the three Vanikoro languages form a subgroup of their own, pointing to a single common ancestor.

1.4 Divergence and convergence among Vanikoro languages

In sum, Teanu, Lovono and Tanema are three ‘aberrant’ Oceanic languages which are genetically closely related to each other, as descendants of a single ancestor. Now as we compare the three modern languages, we may be surprised by a paradox:

- Compared to the situation in most other Oceanic languages, the forms of words in the three languages of Vanikoro tend to differ from each other in ways which can be regarded as quite extreme given their geographic closeness.
- Despite their heterogeneity with regard to word forms, they show perfect isomorphism of their structures.

This paradox will be the main focus of this article.

Throughout this paper, I will refer to two essential components of language, which crosscut the traditional division between lexicon and grammar. On the one hand, I will mention STRUCTURES, referring to the various concepts and semantic categories with which a language divides up semantic space — whether this refers to lexical or to grammatical meaning. On the other hand, each language embodies these categories and concepts into linguistic FORMS, endowed with a specific phonological content.⁹ Two languages can be said to be isomorphic in a particular domain of their system, if they share the same structures or semantic categories, whether or not the forms they use are cognate with each other. For example, the two constructions *I have caught cold* and *J’ai attrapé froid* are perfectly isomorphic, because they express the same event by resorting to exactly identical metaphors and categories.

My observation is that the three Vanikoro languages exhibit a high degree of divergence in their forms, yet still show an extreme isomorphism of their structures. This configuration is illustrated in example (3). As far as the grammar is concerned, the three languages possess parallel structures and word order, to the point that they can all be analysed with a single line of word-to-word glosses. Yet on the other hand, one can equally note the dissimilarity between the actual forms of their words.¹⁰

(3)	TEA	<i>A-ko</i>	<i>u-ka</i>	<i>u-katau</i>	<i>ene ?</i>
	LVN	<i>Nu-pu</i>	<i>ku-ma</i>	<i>ku-ki</i>	<i>ɲane ?</i>
	TNM	<i>Go-po</i>	<i>go-loma</i>	<i>go-ie</i>	<i>nana ?</i>
		2SG:R-say	2SG:IR-come	2SG:IR-follow	1SG
		‘Do you want to come with me?’			

⁹ This contrast STRUCTURES versus FORMS corresponds to what Hjelmslev (1961:52) described as respectively *content form* versus *expression form*; and to what Grace (1981:24) would call *content form* versus *lexification*.

¹⁰ Abbreviations in glosses include: R - realis prefix; IR - irrealis prefix; GEN - General possessive classifier; INDEP - independent pronoun; FOOD - possessive classifier for food possession; HUM - article for human referents.

This observation, whereby languages can be at once homogeneous in structure while dissimilar in forms, has been widely made in the linguistic literature already (see Gumperz 1971; Enfield 2001). However, most of the time, those facts of structural parallelism result historically from contact between genetically diverse languages. To focus on works on the Melanesian area, Thurston (1989, 1994) thus describes the structural similarities between languages of northwestern New Britain, some Papuan and some Oceanic, and Ross (1996, 2001) discusses the influence of Waskia (Papuan) upon Takia (Oceanic), in Papua New Guinea's Madang Province. To this list, one could add discussions of structural parallelism between the English-based pidgins of the Pacific and their Melanesian substrates (Camden 1979; Keesing 1988, 1991; Siegel 2008). In all these cases, whether one compares Oceanic with Papuan or with European languages, the diversity of forms is a given; and what is observed is first and foremost a matter of STRUCTURAL CONVERGENCE — or 'metatypy', to use the term coined by Malcolm Ross (1996, 2001).

But the case of Teanu, Lovono and Tanema raises different issues, because they belong to the same genetic subgroup. For such closely related languages to share the same syntactic structures may partly reflect the mere legacy of their common ancestor, and partly be explained by later contact-induced convergence. The intriguing part here is rather the dissimilarity of forms: it needs to be considered not as a simple given — as was the case for genetically diverse languages — but as the problematic result of historical DIVERGENCE from a common ancestor. This configuration therefore requires specific explanations, beyond the now well-known cases of metatypy.¹¹

Section 2 will discuss the degree of similarity and dissimilarity existing between the lexical forms of Vanikoro languages. Section 3, in turn, will demonstrate their strong structural isomorphism. Section 4 will finally propose a functional hypothesis to account for this linguistic paradox of Vanikoro languages.

2 Similarity and divergence of lexical forms

The linguistic relations between the three Melanesian languages of Vanikoro can thus be summarised in a simple formula: DISSIMILAR FORMS, SIMILAR STRUCTURES. The present §2 will discuss the first of these two dimensions, by assessing the degree of differentiation between the lexicons of Teanu, Lovono, and Tanema.

Even though Teanu, Lovono and Tanema are close genetic relatives (§1.2), the impression that prevails is that of a rather strong dissimilarity of their lexicons. Indeed, contrary to the impression given by Table 1 above, the three languages of Vanikoro are not mutually intelligible. Their dissimilarities are of varying nature, going from more or less regular phonological differences between cognate words, to forms that are simply non-cognate. Overall, while such formal differences are frequent in the Melanesian area, their degree is here rather impressive for languages which are spoken on the same island, and appear to have historically differentiated on this island.

¹¹ A similar blend of formal dissimilarity and structural parallelism can be found in the languages of north Vanuatu (François 2007, forthcoming, in prep.). However, the phenomenon appears to be even more conspicuous in the case of Vanikoro.

2.1 Differences due to phonological change

In some cases, forms which are superficially very dissimilar can in fact be explained by regular processes of sound change. While many correspondences between the three languages of Vanikoro are straightforward and obvious, some are more drastic and can result in little resemblance between the actual forms.¹²

For example, the verbs for ‘sit, stay’ (TEA *te* ~ LVN *lu* ~ TNM *o*) are dissimilar enough to suggest they might be non-cognate. However, one can establish a regular correspondance pattern TEA /t/ ~ LVN /l/ ~ TNM Ø (see sample in Table 3) pointing to a proto-consonant *t (Tryon and Hackman 1983:71).

Table 3: Some lenis reflexes of POc *t

English	Teanu	Lovono	Tanema	POc etymon
‘sit, stay’	<i>te</i>	<i>lu</i>	<i>o</i>	*toka
‘three’	<i>te-te</i>	<i>te-lu</i>	<i>a-o</i>	*tolu
‘sugarcane’	<i>to</i>	<i>lepie</i>	<i>ova</i>	*topu
‘soul, spirit’	<i>ata</i>	<i>ala</i>	<i>ae</i>	*qata
‘ghost, spirit’	<i>tadoe</i>	<i>leñoe</i>	<i>aoe</i>	*qata- [?]
‘do again’	<i>tabo</i>	<i>lebu</i>	<i>abo</i>	
‘carry on shoulders’	<i>tabe</i>	<i>lebe</i>	<i>ebe</i>	
‘impede, protect’	<i>botono</i>	<i>beloŋa</i>	<i>baŋa</i>	
‘unripe, new’	<i>motoe</i>	<i>meloe</i>	<i>maja</i>	*mataq

The three forms for ‘sit, stay’ therefore suggest an etymon *tV. While correspondences are much less obvious regarding vowels, a proto-form with /o/ is a likely origin for a pattern TEA /e/ ~ LVN /u/ ~ TNM /o/; see also the vowels of ‘lie down’ (< POc *qenop) in Table 1, and of ‘three’ in Table 3. The perfectly parallel reflexes for ‘three’ (< *to < POc *tolu) and for ‘sit, stay’ confirm a reconstruction *to — certainly the first syllable of POc *toka ‘stay’. Interestingly, Gaimard (1833, 1834) consistently writes these Tanema words with an *r*, which shows the correspondence pattern was TEA /t/ ~ LVN /l/ ~ TNM /r/ two centuries ago. Thus, he writes *rarou* for ‘three’ (modern *a-o*), and *guidiro* for what he glosses ‘*Asseyez-vous*’; the latter in fact representing **giti-ro* (modern *giti-o*) ‘we_[INCL] are sitting’.

In sum, the regularity of correspondences, when they can be established, makes it possible, quite classically, to detect the cognacy of some forms which would have otherwise seemed unrelated.

Sometimes, the ultimate POc source of a given series is unclear, yet at least one can tentatively draw connections between modern forms, based on synchronic regular sound correspondences. For example, the verb for ‘die’ is *bu* in Teanu and Tanema, and *me* in Lovono, two forms with little in common. However, a regular sound pattern seems once again to emerge from the data: TEA /b/ ~ LVN /m/ ~ TNM /b/ — with a small amount of variation involving voicing or rounding of the consonant (Table 4).

¹² Some of the regular correspondences are given in Ross and Næss (2007).

Table 4: A possible regular correspondence pattern

English	Teanu	Lovono	Tanema	POc etymon
‘die, dead’	<i>bu</i>	<i>me</i>	<i>bu</i>	(*mate) ?
‘sharpen, trim’	<i>bo</i>	<i>me</i>	<i>bo</i>	
‘1exc:dual pronoun’	<i>keba</i>	<i>gema</i>	<i>gabe</i>	*kama[m]i
‘k.o. basket’	<i>iunubo</i>	<i>nunumie</i>	<i>nuba</i>	
‘tobacco’	<i>nabene</i>	<i>nakamene</i>	<i>nabwane</i>	
‘blood’	<i>abo</i>	<i>amwale</i>	<i>aba</i>	
‘Areca catechu’	<i>buioe</i>	<i>namwe</i>	<i>buia</i>	*buaq
‘Reflexive-Reciprocal’	<i>ñepe</i>	<i>ñeme</i>	<i>be</i>	
‘fishing net’	<i>pele</i>	<i>menele</i>	<i>benala</i>	*kup ^w ena ?

This suggests the three forms for ‘die’ may be cognate, despite their present dissimilarity. In this case, because no known POc reconstruction (including *mate) provides any satisfying etymon, the cognacy judgment rests on purely synchronic data.

Table 5: Some potential, but dubious, cognate sets

English	Teanu	Lovono	Tanema	POc etymon
‘man, person’	<i>mwaliko</i>	<i>lamuka</i>	<i>anuka</i>	*m ^w aqane ?
‘women’	<i>viñevi</i>	<i>veñime</i>	?	*pine ?
‘canoe, ship’	<i>kuo</i>	<i>nawe</i>	<i>goia</i>	*waga(ŋ) ?
‘sleep’	<i>mokoiu</i>	<i>mepu</i>	<i>matou</i>	*maturur ?
‘red’	<i>moloe</i>	<i>wamoene</i>	<i>manobeila</i>	*meraq ?
‘randomly; in vain’	<i>moli</i>	<i>moli</i>	<i>mano</i>	
‘help; with’	<i>samame</i>	<i>emeio</i>	<i>avaio</i>	
‘where?’	<i>vele</i>	<i>mane</i>	<i>vane</i>	
‘I, 1sg pronoun’	<i>ene</i>	<i>ŋane</i>	<i>nana</i>	
‘sink’	<i>metelu</i>	<i>mwelesu</i>	<i>madilo</i>	
‘perhaps’	<i>bwara</i>	<i>bweti</i>	<i>buru</i>	
‘go down’	<i>abu</i>	<i>pwo</i>	<i>kabu</i>	
‘good’	<i>wako</i>	<i>vakane</i>	<i>apika</i>	
‘seize, hold’	<i>labu</i>	<i>lo</i>	<i>nou</i>	
‘rejoice’	<i>pei</i>	<i>pwadi</i>	<i>pae</i>	
‘stone oven’	<i>awene</i>	<i>epene</i>	<i>pavene</i>	

Occasionally, the modern forms display little more than a vague ‘family resemblance’. That is, the modern lexical forms are possibly cognate, yet their phonemes enter no regular correspondence pattern, in such a way that one could only explain their cognacy by resorting to *ad hoc* etymological hypotheses. Table 5 provides a sample of such potential, but irregular and dubious, cognate sets, which would require closer scrutiny in the future.

2.2 Lexical replacement

Finally, it also often happens that the three languages have forms that are distinct, and almost certainly not cognate for the same meaning. A sample of such cases is given in Table 6.

Table 6: Some non-cognate sets resulting from lexical innovation

English	Teanu	Lovono	Tanema
'thing'	<i>ɲatene</i>	<i>vesemele</i>	<i>vamora</i>
'know'	<i>ovei</i>	<i>lonei</i>	<i>wo</i>
'come'	<i>ka</i>	<i>mage</i>	<i>loma</i>
'lie, deceive'	<i>tomoli</i>	<i>ñaine</i>	<i>role</i>
'see'	<i>romo</i>	<i>eti</i>	<i>runi</i>
'quickly'	<i>kiane</i>	<i>segei</i>	<i>gamoi</i>
'bad'	<i>tamwaliko</i>	<i>visale</i>	<i>vae</i>
'big (PLUR)'	<i>wopine</i>	<i>evala</i>	<i>bwau</i>
'broken'	<i>mamakoe</i>	<i>seli</i>	<i>vave</i>
'remote'	<i>somu</i>	<i>akaole</i>	<i>mosomu</i>
'down, below'	<i>puo</i>	<i>lenu</i>	<i>ese</i>
'what?'	<i>(ɲan)ae</i>	<i>ese</i>	<i>sive</i>
'do what, do how?'	<i>(mi)kae</i>	<i>ñese</i>	<i>jive</i>
'be why?'	<i>ve</i>	<i>wo</i>	<i>ja</i>
'another, an'	<i>iote</i>	<i>leka</i>	<i>keo</i>
'one; same'	<i>iune</i>	<i>tilioko</i>	<i>omwano</i>
'play; wander'	<i>moloe</i>	<i>telu</i>	<i>lumota</i>
'neck; mind'	<i>awa</i>	<i>warene</i>	<i>vasare</i>
'bush'	<i>ɲogoro</i>	<i>atere</i>	<i>arara</i>
'cabbage'	<i>tebo</i>	<i>lamware</i>	<i>some</i>
'rat'	<i>uvilo</i>	<i>katone</i>	<i>ivala</i>
'be ripe'	<i>ako</i>	<i>wi</i>	<i>kou</i>
'guts'	<i>bea</i>	<i>lale</i>	<i>lebwe</i>
'year'	<i>ebieve</i>	<i>verue</i>	<i>rove</i>
'chicken'	<i>kulevelu</i>	<i>kio</i>	<i>tokila</i>
'leaf'	<i>uie</i>	<i>nugia</i>	<i>lele</i>

In their survey of Solomon Island languages, Tryon and Hackman (1983:481) give the following cognate percentages for the three languages of Vanikoro, based on a modified version of Swadesh's basic vocabulary list (200 words):

- Teanu–Lovono 57.3 %
- Teanu–Tanema 51.1 %
- Lovono–Tanema 54.8 %

These figures point to a relatively high level of lexical dissimilarity. They surpass similar counts made in most other parts of island Melanesia. For example, the Torres and Banks Is

of north Vanuatu are another area where lexical replacement has been intense (François, in prep.). However, in order to find figures as low as those on Vanikoro, one has to pick languages which are geographically spread apart; e.g. 44.5 % between Hiw (Torres) and Lakon (Gaua, south Banks). Should one consider languages spoken on a single island, the widest gap one can find there is between Lakon and Dorig, with 61.5 % shared vocabulary (Tryon 1976:95).

Even more instructive is the comparison with other language families in the world. To take just one example, rates of shared vocabulary amongst Germanic languages do not go below 53.6% (Dyen, Kruskal and Black 1992). In other words, the three languages of Vanikoro have managed to achieve, within the limited space of a single island, more lexical diversity than the whole Germanic subgroup of Indo-European. Such an intense differentiation deserves to be acknowledged, and discussed (see §4).

The strong formal dissimilarity that prevails between the three languages of Vanikoro naturally results in sentences where the phonological form of words, whether lexical or grammatical, can show a high degree of dissimilarity:

- (4) TEA *Pi-te* *ne sekele* *iupa,* *pi-wowo* *uo.*
 LVN *Nupe-lu* *ne amenoja* *iemitore,* *nupe-ŋoa* *upie.*
 TNM *Tei-o* *ini vasaŋola* *akegamuto,* *ti-oa* *uva.*
 1EX.PL:R-stay in garden our 1EX.PL:R-plant yam
 ‘We were in our garden, we’ve been planting yams.’
- (5) TEA *Somu* *tamwase* *tae,* *vitoko* *takoie* *ne.*
 LVN *Akaole* *visalewabeu* *taie,* *vateoko* *ŋate* *ida.*
 TNM *Mosomu* *vaepamabo* *eia,* *vatako* *eto* *kana.*
 remote very NEG close uphill here
 ‘It is not very far; it’s close to here, up this way.’

The impression of formal heterogeneity may be due partly to sound change affecting words that are in fact cognate (e.g. ‘yam’, ‘stay’, ‘plant’); and partly to lexical replacement (e.g. forms for ‘very’, ‘remote’, ‘uphill’) or morphological change (e.g. forms of personal pronouns).

What is perhaps more puzzling is the contrast between, on the one hand, this formal dissimilarity, and on the other hand, the perfect parallelism existing between these languages’ structural properties. This is the topic of the next section.

3 Structural isomorphism

The push towards linguistic differentiation has affected the phonological forms of words, yet evidently had little impact on grammatical and semantic structures. Even as their lexicons were diverging from each other, the three languages of Vanikoro have maintained a strong STRUCTURAL ISOMORPHISM — whether in syntax, phraseology or organisation of meaning. This is reflected, for instance, by the ability of translating word-for-word any sentence from one language to another, following the same word order and the same morphological and lexical categories — as in examples (3) to (5).

The present section will illustrate this strong structural parallelism using a few conspicuous examples from various aspects of the grammar, and will end with a tentative account of the linguistic history of Vanikoro.

3.1 Morphosyntax

3.1.1 Syntax of the clause

Teanu, Lovono and Tanema are parallel in all aspects of their syntax. This includes all properties related to word order, whether the basic SVO clause order, or other properties, use of prepositions, post-nominal placement of adjectives and possessors, clause-final position of the negation as in (5).

The three languages display the same organisation in parts of speech, including a sharp divide between nouns and verbs, and a distinctive category of adjectives. While all verbs are obligatorily prefixed for subject and mood (§3.1.2), adjectives are unprefixed. Languages even agree on whether they treat a specific notion as an adjective or as a verb, as evidenced by the parallel presence vs absence of the subject prefix in the two predicates of (6).

- (6) TEA *Udo ponu, boro* *we i-ako ?*
 LVN *Puŋa pae, bware* *we i-wi ?*
 TNM *Uda pade, betika* *we i-kou ?*
 banana that black/unripe or 3SG:R-be.ripe
 ‘Those bananas, are they green_[ADJ] or ripe_[VERB]?’

Likewise, all syntactic properties of the clause are reflected alike across the three languages. They all lack noun articles, case markers, verb transitivisers or applicatives, and causative affixes. They make use of a reflexive marker (Table 4), which is also a reciprocal and an emphatic marker. They all resort frequently to core-layer verb serialisation, with exactly parallel phrasing; see examples (3), (7) and (10). They possess exactly parallel Tense-Aspect-Mood categories, and so on.

3.1.2 Pronouns and TAM marking

The paradigms of personal pronouns are also organised in parallel ways. Like most Oceanic languages, those of Vanikoro distinguish between exclusive and inclusive ‘we’; and they show three numbers: singular, dual, plural.

Table 7 shows the eleven independent pronouns for the three languages. These show a reasonable degree of similarity.

Table 7: Personal independent pronouns

	Teanu	Lovono	Tanema
1sg	ene	ŋane	nana
2sg	eo	ago	go
3sg	ini	ŋani	nini
1in:du	kia	gita	gie
1ex:du	keba	gema	gabe
2du	kela	gamila	gamile
3du	da	dea	delalu
1in:pl	kiapa	gitu	geto
1ex:pl	kupa	gamitu	gamuto
2pl	kaipa	gaipa	gamito
3pl	dapa	detu	dato

Besides these free pronouns, these languages also possess¹³ a double set of mood-marked subject prefixes for verbs (one for realis, one for irrealis); see Table 8.

Table 8: Verbal prefixes for subjects

	Teanu		Lovono		Tanema		
	Realis	Irrealis	Realis	Irrealis	Realis	Irrealis	
1sg	ni-	ne-	ni-	ka-	ne/i-	na-	
2sg	a-	u-	nu-	ku-	go/i-	go-	
3sg	i-	i-	i-	ki-	i-	i-	
1in:du	la(i)-	la(i)-	la(i)-	sa-	de-	ja-	}
1ex:du	ba(i)-	ba(i)-	(nu)ba-	ba(i)-	ba(i)-	ba(i)-	
2du	ba(i)-	ba(i)-	(nu)ba-	ba(i)-	ba(i)-	ba(i)-	
3du	la(i)-	la(i)-	la(i)-	sa-	de-	ja-	}
1in:pl	li-	le-	le(pe)-	kape-	le/i-, giti-	la-	
1ex:pl	pi-	pe-	nupe-	pe-	te/i-	tu-	
2pl	pi-	pe-	nupe-	pe-	te/i-	tu-	}
3pl	li-	le-	le(pe)-	se(pe)-	le/i-	la-	

Beyond their general family resemblance, the subject prefixes shown in Table 8 show a certain amount of formal variety, see especially the 2sg, or the plural forms. Yet once again, this formal diversity goes along with a strong structural isomorphism. In particular, all languages display the same two morphological mergers on non-singular pronouns: merger of 1st inclusive and 3rd person (with a couple of exceptions) on the one hand, and merger of 1st exclusive and 2nd person on the other. This morphological pattern is specific to the three Vanikoro languages, and is not found in neighbouring Utupua (Tryon 1994:631). This may therefore constitute an important shared innovation diagnostic of a Vanikoro subgroup (see §1.3.2).¹⁴

Finally, our three languages add to the set of free pronouns another personal category, namely 3rd PLURAL INDEFINITE. This category has a special form as a free pronoun — TEA *idi* ~ LVN *nili* ~ TNM *deli*, which may be glossed ‘people’ (cf. French *on*). When this free pronoun is the subject, the agreement marker on the verb will be an ordinary 3rd plural prefix.

3.1.3 Possessive classes

The three languages also agree in the morphosyntax of possession. They all encode inalienable possession identically, by juxtaposing the possessed noun and its possessor. In the absence of possessive suffixes, inalienable possessors are encoded with the

¹³ Instead of being coded by dedicated suffixes as in POc, objects and inalienable possessors are expressed by independent pronouns, which form distinct phonological words. The only exception to this principle is the Lovono suffix *-ŋo* for 2sg objects and possessors, which is distinct from the free pronoun *ago*, see Table 13 below.

¹⁴ Pronoun systems often provide crucial diagnostic evidence in subgrouping research (Ross 2005).

independent personal pronoun: e.g. TEA *awa ini* ~ LVN *warene nyani* ~ TNM *vasare nini* /throat 3sg:INDEP/ ‘his throat’. Semantically, inalienable possession covers most body parts, plus a handful of intimate belongings, e.g. TEA *bete ene* ~ LVN *bele nyane* ~ TNM *be nana* /mat 1sg:INDEP/ ‘my bedmat’.

A possessive classifier is required for alienable types of possession. Four possessive categories can be distinguished (I indicate in square brackets the Teanu form of the classifier for 1sg possessor): FOOD [*enaka*]; DRINK [*me ene*]; KINSHIP [*one*]; GENERAL possession [*enone*], used as a default.

The category of FOOD possession is larger than its label suggests. First, it covers food (including items only chewed, like areca nut and betel leaf) as well as drink, thus overlapping with the dedicated DRINK classifier. Second, it is required for most tools (‘knife’, ‘adze’, ‘spear’, ‘hook’, ‘box’...), plus the generic term ‘belongings’. Third, it is used for ‘language’ and ‘custom’.

- (7) TEA *U-labu nyatene enaka u-lui ne mwoe enone.*
 LVN *Ku-lo vesemele aña ku-lawoi ne moe iaña.*
 TNM *Go-nou vamora ae go-lao ini nalama ie.*
 2SG:IR-*hold* thing FOOD:1SG 2SG:IR-*take.away* in house GEN:1SG
 ‘Get my_[FOOD] belongings and take them to my_[GENERAL] house.’

- (8) TEA *Dapa iakapa kape le-mui piene akapa.*
 LVN *Detu iegitore gape se-moi mwamwane agitore.*
 TNM *Dato egeto mota la-muo puiene egeto.*
 PL:HUM GEN:1INC:PL FUT 3PL:IR-*not.know* speech FOOD:1INC:PL
 ‘Our_[GENERAL] people are going to forget our_[FOOD] language.’

Once again, the three languages of Vanikoro agree perfectly on the semantic content of their formal categories.

3.1.4 Space directionals

Another domain where Teanu, Lovono and Tanema share identical structures is the system of space reference. Even though their adverbial directionals show impressive formal diversity (Table 9), their functional properties are parallel: they all resort to the ‘in’–‘out’ contrast to encode the sea–land axis, and use the ‘up’–‘down’ pair to encode a fixed cardinal axis, oriented towards southeast. While this system is attested elsewhere among Oceanic languages, it is distinct from the one reconstructed for POc (François 2004).

Table 9: Space directionals in their local and geocentric uses

Local use	Geocentric use	Teanu	Lovono	Tanema
‘in’	‘inland’	<i>takoie</i>	<i>ηate</i>	<i>eto</i>
‘out’	‘seaward’	<i>tetake</i>	<i>mwaroa</i>	<i>emo</i>
‘up’	‘toward SE’	<i>tev’ iu</i>	<i>ηau</i>	<i>iu</i>
‘down’	‘toward NW’	<i>tev’ tawo</i>	<i>lenu</i>	<i>ese</i>

3.2 Lexicon

The structural isomorphism so characteristic of Vanikoro languages relates not only to the morphosyntax, but also to the semantic organisation of the lexicon. I will mention successively two types of subdomain where this parallelism can be observed: the lexicon proper, and the phraseology.

3.2.1 Lexical semantics

We have seen that Teanu, Lovono and Tanema provide each grammatical category (possessive classifiers, space directionals ...) with essentially the same semantic outline. The same can be said of lexical items and their meaning: when two quite distinct meanings are ‘colexified’ in one language — i.e. are expressed by the same lexical form (François 2008b) — the same pattern of colexification will almost certainly be found in the two other languages. Setting aside cases of polysemy which are shared by all or most Oceanic languages (e.g. ‘hear’–‘feel’ ...), some of the most distinctive examples of colexification are shown in Table 10.

Table 10: Most colexification patterns are shared across Vanikoro languages

Sense 1	Sense 2	Teanu	Lovono	Tanema
‘one, single’	‘the same’	<i>iune</i>	<i>tilu ~ tilioko</i>	<i>omwano</i>
Indefinite SG	‘another’	<i>iote</i>	<i>leka</i>	<i>keo</i>
Indefinite PL	‘others’	<i>kula</i>	<i>kule</i>	<i>kule</i>
‘all’	‘many’	<i>abia</i>	<i>maraja</i>	<i>abia</i>
‘bird’	‘friend’	<i>menuko</i>	<i>menuka</i>	<i>manuke</i>
‘light (adj.)’	‘dry’	<i>mimione</i>	<i>mimiane</i>	<i>mamiene</i>
‘black’	‘unripe’	<i>boro</i>	<i>bware</i>	<i>betika</i>

In the domain of compounding, a special case can be made regarding three adjectives, represented in Table 11. Despite their formal differences, they appear to be everywhere analysable in the same way, as if the result of calquing. Most forms are synchronically transparent, and based on the noun ‘name’. As for TEA *ɲasune*, it can be analysed as historically a combination of *(e)ɲa* ‘name’ and *iune* ‘one, the same’.

Table 11: Lexical connection between noun ‘name’ and three adjectives

English	Teanu	Lovono	Tanema	Literally
‘name’	<i>eɲa</i>	<i>neɲe</i>	<i>niɲe</i>	
‘identical’	<i>ɲasune</i>	<i>neɲe-tilu</i>	<i>niɲe-omwano</i>	‘name-one’
‘different’	<i>eɲa-iote</i>	<i>neɲe-leka</i>	<i>niɲe-keo</i>	‘name-other’
‘various’	<i>eɲa-eɲa</i>	<i>neɲe-neɲe</i>	<i>niɲe-niɲe</i>	‘name-name’

Likewise, each of the three languages derives its intensifier ‘very much, too much’ — see ex.(5) — from its adjective ‘bad’.¹⁵ This connection is especially noteworthy as it involves distinct roots in each language (Table 12). The second element in these compound forms is obscure.

Table 12: Lexical connection between adjective ‘bad’ and intensifier

English	Teanu	Lovono	Tanema
‘bad’	<i>tamwaliko</i>	<i>visale</i>	<i>vae</i>
Intensifier	<i>tamwa(liko)se</i>	<i>visale-wabeu</i>	<i>vae-pamabo</i>

3.2.2 Phraseology

The structural isomorphism between the three languages of Vanikoro is equally obvious from their phraseology, i.e. the routinised way in which they connect words together. I will only mention here a couple of original cases.

The three languages have an inalienable noun for ‘body’, to which they attach a variety of meanings, including ‘genuine, true’¹⁶ and ‘beautiful’, see Table 13.

Table 13: The polysemy of the noun ‘body’

TEA	<i>ebele</i>	<i>eo</i>	<i>ebele</i>	<i>piene</i>	<i>ebele</i>	<i>kuo</i>	<i>ebel’</i>	<i>ini</i>
LVN	<i>nebele</i>	<i>-ŋo</i>	<i>nebele</i>	<i>mwamwane</i>	<i>nebele</i>	<i>nawe</i>	<i>nebele</i>	<i>ŋani</i>
TNM	<i>nibela</i>	<i>go</i>	<i>nibela</i>	<i>puiene</i>	<i>nibela</i>	<i>goia</i>	<i>nibela</i>	<i>nini</i>
	body	2sg	body	speech	body	canoe	body	3sg
	‘your body’		‘true words, truth’		‘canoe hull’ ~		‘his/her/its body’ ~	
					‘beautiful canoe’		‘Wonderful!’	

The neck or throat evidently constitutes, in Vanikoro, the seat of emotions and feelings.¹⁷ The corresponding noun is found in a variety of formulas:

- ‘I’m angry’ is literally ‘*My throat is burning*’
- ‘I’m sad’ is literally ‘*My throat is blocked*’

Our three Vanikoro languages are strictly parallel in all these formulations. This is a fact of structural isomorphism, or calquing, as the words for ‘throat’ do not appear to be cognate across languages:

¹⁵ The connection is also attested in English (*I want it badly*), and closer to Vanikoro, in the Torres languages of Vanuatu: e.g. Lo-Toga *na luwō hia*, lit. ‘it’s big bad’ = ‘it’s too big’ (François, pers. data).

¹⁶ Interestingly, the languages of northern Vanuatu share the same colexification pattern between ‘body’ and ‘true’. François (2005:501) thus proposes to reconstruct, for the common ancestor of north Vanuatu languages, a protoform *tur[i,u](ʔi) ‘body, trunk; the real, main, very X; really’.

¹⁷ Osmond (2007) reports similar metaphors of emotions located in the larynx, for languages of the Southeast Solomons, as well as for the languages of the Trobriand Islands (after Malinowski 1922:408).

- (9) TEA *Awa kupa i-su.*
 LVN *Warene gamitu i-tu.*
 TNM *Vasare gamuto i-to.*
 throat 1EX:PL:INDEP 3SG:R-blocked
 [lit. ‘Our throats are blocked.’] ‘We’re sad ~ We’re sorry.’

The sense ‘like’/‘want’ is expressed by an unusual formula using a verb ‘hit’, taking the ‘throat’ as its subject:

- ‘I like/want this’ is literally ‘*My throat is hitting this.*’

The sentence becomes even more unusual when it is followed by an object clause (‘want to do’), because it then involves a complementiser which is literally a verb meaning ‘say’.¹⁸ The subject of ‘say’ is normally the ‘throat’ itself (hence 3sg agreement), but occasionally it agrees syntactically with the throat’s possessor:

- ‘I want to [sleep]’ is literally ‘*My throat is hitting IT SAYS I [sleep] ...*’
 or ‘*My throat is hitting I SAY I [sleep] ...*’
 which is often shortened to ‘*My throat I SAY I [sleep] ...*’

- (10) TEA *Awa ene (i-viaene) ni-ko ne-mokoiu.*
 LVN *Warene ηane (i-piaine) ni-pu ka-mepu.*
 TNM *Vasare nana (i-vini) ni-po na-matou.*
 throat 1SG:INDEP 3SG:R-hit 1SG:R-say 1SG:IR-sleep
 ‘I want to sleep.’

In this case just as in all other contexts, the three languages can be translated literally, morpheme-by-morpheme, with no loss in idiomaticity or change in meaning. All one has to do is keep the structural — grammatical and lexical — boxes, and swap their phonological contents.

4 Addressing the paradox

In sum, the three languages of Vanikoro can be characterised by two contradictory properties. On the one hand, their fundamental genetic relatedness is blurred by a high degree of dissimilarity in the phonological forms of words, whether in the lexicon or in the morphology. But on the other hand, their grammatical categories and semantic structures show no equivalent to this formal diversity: instead, the three languages reveal perfect isomorphism, in each and every corner of their system. To paraphrase a formula by Sasse (see fn.19 below), they could ultimately be described as ‘*a single language with different vocabularies*’.

The question arises of what historical scenario would best explain this paradox, where divergence goes along with convergence. A simple explanation that comes to mind when accounting for the lexical diversification of cognate languages, might focus on the physical

¹⁸ The grammaticalisation of a verb of saying into a complementiser is typologically common (Heine and Kuteva 2002; Chappell 2008). To take an Oceanic example, the verb ‘say’ in Araki, Vanuatu (François 2002), has exactly the same properties as in Vanikoro languages, including the persistence of a fully verbal morphology even when used as a complementiser.

separation between language communities. The absence, still today, of any land path relating villages across Vanikoro island, and the stories of ongoing fierce territorial fights between its three tribes, would then be understood as genuine evidence for geographical or social isolation, and thus as a possible key for the high degree of formal divergence between Teanu, Lovono and Tanema. However, several facts seem to contradict this diagnostic. First, the relatively small size of the island is at odds with the notion of a neat separation between the three tribes. And more crucially, their extreme degree of structural isomorphism is likely to reflect not only cases of shared retentions from a common ancestor, but also later linguistic convergence induced by language contact. In other words, the explanation resorting to the mere physical separation between communities does not tell the whole story.

The solution to the puzzle will probably have to be found not in the factual features of geography, but in the more subtle dimension of sociolinguistic behaviour. Indeed, a conspicuous characteristic of cultures in certain parts of Melanesia — in comparison, for example, with the Polynesian world (see Pawley 1981) — seems to be a social preference for small-scale social communities with no marked hierarchy between them, as well as a strong emphasis put on whichever anthropological or linguistic features may *differ* from one community to the other. Heterogeneity between villages or village groups tends to be socially valued as a way to construct a world of diversity, where each community is endowed with its own identity. In this framework, a local innovation in cultural and linguistic forms will tend to be perceived, and eventually retained, as emblematic of a specific group. Over time, this behaviour favours the emergence of cultural and linguistic divergence between erstwhile homogeneous communities. Interestingly, some language groups can be said to have only gone down this track to the point when the languages began to lose mutual intelligibility; but what is conspicuous in the case of Vanikoro languages, is that they seem to have pushed the process of differentiation far beyond that point, as though they were to keep diverging for ever.

In order to account for similar facts in other parts of Papua New Guinea, Thurston (1989), and later Ross (1996; 2001:155), have used the term ESOTEROGENY:

Esoterogeny is a process that adds structural complexity to a language and makes it more efficient as a medium of communication among people of the same social group, while making it more difficult for outsiders to learn to speak well. (Thurston 1989)

Esoterogeny arises through a group's desire for exclusiveness. (Ross 1996:184)

If the members of a community have few ties with other communities and their emblematic lect is not usually known to outsiders, then they may use it as an 'in-group' code, an 'esoteric' lect from which outsiders are consciously excluded. Innovations leading to increased complexity and to differences from neighbouring lects will be favoured. (Ross 1997:239)

One could probably discuss the degree to which such sociolinguistic processes are 'conscious', and also how they interfere with motivations of various kinds (semantic, structural, pragmatic) in bringing about change. This being said, one can probably accept the general idea behind Thurston's concept, that language differentiation in Melanesia, far from being just an accident of geographical isolation, is largely influenced by a certain social attitude whereby each group tends to produce — whether consciously or not — its own distinctive speech tradition.

Now, while this hypothesis may help explain the high amount of lexical innovation and formal divergence that took place between Vanikoro languages, it seems at odds with the remarkable stability that we've observed among their structures. I would suggest this mismatch can be explained by the different nature of the linguistic components involved here. For one thing, the phonological form of the words (Saussure's 'signifiant', Grace's 'lexification'), whether lexical or grammatical, is the component most salient and conspicuous to the speakers' conscience, and therefore most likely to be preempted by motivations based on social emblematicity. Conversely, the structural and semantic dimension of language (Saussure's 'signifié', Grace's 'content form') would fall out of reach of the speakers' immediate linguistic awareness, in a way that would make it exempt of the sociolinguistic force of *esoterogeny*. Instead, structures tend to obey a totally contrary force, typical of language-contact situations,¹⁹ that leads them to diffuse and converge: this is when multilingual speakers feel the pressure 'towards word-for-word translatable codes' (Gumperz 1971:270). The structural isomorphism that can be observed today among Vanikoro languages has the considerable advantage, for the bilingual speaker, of reducing any translation loss, thereby increasing the efficiency of cross-linguistic communication, and facilitating the cognitive processing of speech.

An important corollary of this whole reasoning is the necessity to distinguish two different components of language, because their evolution through history can follow quite distinct paths:

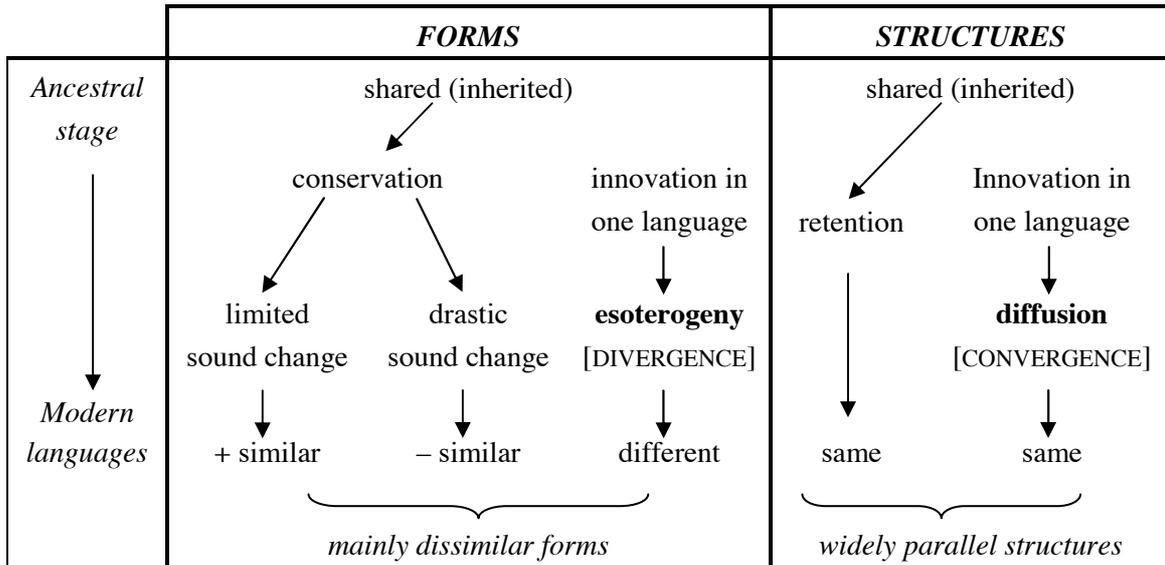
The two components of language — content form and lexification — (...) evolve independently because (...) they are responding to different selective pressures, and those selective pressures are different because the functions of the two components are different. (...) It is the lexification on which the emblematic burden ultimately falls. (Grace 1981:30)

Vanikoro illustrates an extreme case of this possible mismatch between the two components of language.

Table 14 summarises a possible scenario about the sequence of developments in the evolution of the Vanikoro languages.

¹⁹ Among many other references, see in particular Malcolm Ross' (1996, 1997, 2001) concept of *metatypy*, i.e. the typological alignment of one language to the structures of a neighbouring language, through linguistic contact. Ross (2001:149) also cites this statement by Sasse (1985): 'With advanced language contact, there arises the tendency to develop a single language with different vocabularies.'

Table 14: Different forms, shared structures among Vanikoro languages: a historical scenario



5 Conclusion

The comparison of Teanu, Lovono and Tanema reveals the intricacies of the island’s local history. The strong isomorphism found between the structures of these languages betrays their remote common ancestry, as much as it points to a history of intense language contact which the three tribes, *volens volens*, have lived through over the centuries. On the other hand, the actual word forms found in their vocabularies and morphology have tended to follow a powerful tendency towards diversification, in accordance with the speakers’ tacit perceptions that the three communities, often caught in conflict and territorial hostilities, should sound and feel to be distinct social groups.

Overall, the paradox observed among the three modern languages of Vanikoro — *dissimilar forms, similar structures* — results from the interplay between these two contradictory forces: a socially driven push to increase language differences versus a functionally grounded tendency to minimise them.

References

Camden, William. 1979. Parallels in structure and lexicon and syntax between New Hebrides Bislama and the South Santo language spoken at Tangoa. In *Papers in Pidgin and Creole linguistics, No.2*, 51–117. Canberra: Pacific Linguistics.

Chappell, Hilary. 2008. Variation in the grammaticalization of complementizers from *verba dicendi* in Sinitic languages. *Linguistic Typology* 12.45–98.

Dillon, Peter. 1830. *Voyage aux îles de la mer du Sud, en 1827 et 1828, et Relation de la découverte du sort de La Pérouse. Par le Capitaine Peter Dillon*. Paris: Pillet Aîné.

Dumont d’Urville, Jules Sébastien César. 1830–1834. *Voyage de la corvette l’Astrolabe, exécuté par ordre du Roi, pendant les années 1826–1827–1828–1829, sous le commandement de M.J. Dumont d’Urville, Capitaine de vaisseau*. 23 volumes. Paris: J. Tastu.

- Dyen, Isidore, Joseph B. Kruskal, and Paul Black. 1992. *An Indo-European classification: a lexicostatistical experiment*. Transactions of the American Philosophical Society, vol. 82(5). Philadelphia: The American Philosophical Society.
- Enfield, Nicholas. 2001. On genetic and areal linguistics in mainland Southeast Asia: parallel polyfunctionality of 'acquire'. In Alexandra Aikhenvald and R.M.W. Dixon, eds *Areal diffusion and genetic inheritance: problems in comparative linguistics*, 255–290. Oxford: Oxford University Press.
- Firth, Raymond. 1985. *Tikopia–English dictionary – Taranga fakatikopia ma taranga fakainglisi*. Auckland: Auckland University Press.
- François, Alexandre. 2002. *Araki. A disappearing language of Vanuatu*. Canberra: Pacific Linguistics.
- — 2004. Reconstructing the geocentric system of Proto Oceanic. *Oceanic Linguistics* 439(1).1–32.
- — 2005. Unraveling the history of the vowels of seventeen northern Vanuatu languages. *Oceanic Linguistics* 449(2).443–504.
- — 2006. Are Vanikoro languages really Austronesian? Paper presented at the Second conference on Austronesian Languages and Linguistics. Oxford, UK.
- — 2007. Parallel meanings, divergent forms in the North Vanuatu Sprachbund: diffusion or genetic inheritance? Paper presented at 7th International Congress of the Association of Linguistic Typology. Paris, France.
- — 2008a. Mystère des langues, magie des légendes. In Association Salomon, ed. *Le mystère Lapérouse ou le rêve inachevé d'un roi*, 230–233. Paris: de Conti – Musée national de la Marine.
- — 2008b. Semantic maps and the typology of colexification: intertwining polysemous networks across languages. In M. Vanhove, ed. *From polysemy to semantic change: towards a typology of lexical semantic associations*, 163–215. Amsterdam/Philadelphia: John Benjamins.
- — forthcoming. Des valeurs en héritage: les isomorphismes sémantiques et la reconstruction des langues. In I. Choi-Jonin, M. Duval and O. Soutet, eds *Typologie et comparatisme. Hommages offerts à Alain Lemaréchal*. Orbis-Supplementa 29. Louvain: Peeters.
- — in preparation. Language ecology and change in north Vanuatu: a tale of divergence and convergence.
- Gaimard, Joseph Paul. 1833. Extraits du Journal de M. Gaimard. In J. Dumont d'Urville, *Voyage de la corvette l'Astrolabe, exécuté par ordre du Roi, pendant les années 1826–1827–1828–1829, sous le commandement de M.J. Dumont d'Urville, Capitaine de vaisseau – Histoire du voyage*, vol. 5, 322–353. Paris: J. Tastu, éditeur.
- — 1834. Vocabulaires des idiomes des habitans de Vanikoro. In J. Dumont d'Urville, *Voyage de découvertes de l'Astrolabe, exécuté par ordre du Roi, pendant les années 1826–1827–1828–1829, sous le commandement de M.J. Dumont d'Urville, Capitaine de vaisseau – Philologie*, vol. 1, 165–174. Paris: Ministère de la Marine.
- Grace, George. 1981. *An essay on language*. Columbia, SC: Hornbeam Press.

- — 1990. The ‘aberrant’ (vs. ‘exemplary’) Melanesian languages. In Philip Baldi, ed. *Linguistic change and reconstruction methodology*, vol. 45, 155–173. Berlin: Mouton de Gruyter.
- Gumperz, John. 1971. Convergence and creolization: a case from the Indo-Aryan/Dravidian border in India. In *Language in social groups: essays by John J. Gumperz*, 251–273. Stanford: Stanford University Press.
- Heine, Bernd and Tania Kuteva. 2002. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.
- Hjelmslev, Louis. 1961 [1943]. *Prolegomena to a theory of language*. Madison: University of Wisconsin Press.
- Ivens, Walter G. 1918. *Dictionary and grammar of the language of Sa’a and Ulawa, Solomon Islands*. Washington: Carnegie Institution of Washington.
- Keesing, Roger M. 1988. *Melanesian Pidgin and the Oceanic substrate*. Stanford: Stanford University Press.
- — 1991. Substrates, calquing and grammaticalization in Melanesian Pidgin. In E.C. Traugott and B. Heine, eds *Approaches to grammaticalization*, vol. 1, 315–342. Amsterdam/Philadelphia: John Benjamins.
- Lynch, John. 2001. *The linguistic history of Southern Vanuatu*. Canberra: Pacific Linguistics.
- Malinowski, Bronislaw, 1922. *Argonauts of the western Pacific*. London: Routledge and Kegan Paul.
- Osmond, Meredith. 2007. Describing emotions and enduring dispositions in Oceanic languages. Paper presented at 7th International Conference on Oceanic Linguistics. Nouméa, New Caledonia.
- Pawley, Andrew. 1981. Melanesian diversity and Polynesian homogeneity: a unified explanation for language. In J. Hollyman and A. Pawley, eds *Essays in Pacific languages and cultures in honour of Bruce Biggs*, 269–310. Auckland: Linguistic Society of New Zealand.
- — 2006. Explaining the aberrant Austronesian languages of southeast Melanesia: 150 years of debate. *Journal of the Polynesian Society* 115.213–256.
- — 2007. Was there early Lapita settlement of the Solomon Islands? Bringing linguistic evidence to an archaeological debate. Paper presented at 7th International Conference on Oceanic Linguistics. Nouméa, New Caledonia.
- Ross, Malcolm D. 1988. *Proto Oceanic and the Austronesian languages of western Melanesia*. Canberra: Pacific Linguistics.
- — 1996. Contact-induced change and the comparative method: cases from Papua New Guinea. In Mark Durie and Malcolm Ross, eds *The comparative method reviewed: regularity and irregularity in language change*, 180–217. Oxford: Oxford University Press.
- — 1997. Social networks and kinds of speech-community event. In Roger Blench and Matthew Spriggs, eds *Archaeology and language 1: theoretical and methodological orientations*, 209–261. London: Routledge.

- — 2001. Contact-induced change in Oceanic languages in north-west Melanesia. In Alexandra Aikhenvald and R.M.W. Dixon, eds *Areal diffusion and genetic inheritance: problems in comparative linguistics*, 134–166. Oxford: Oxford University Press.
- — 2005. Pronouns as a preliminary diagnostic for grouping Papuan languages. In A. Pawley, R. Attenborough, R. Hide and J. Golson, eds *Papuan pasts: cultural, linguistic and biological histories of Papuan-speaking peoples*, 15–66. Canberra: Pacific Linguistics.
- Ross, Malcolm D. and Åshild Næss. 2007. An Oceanic origin for Äiwoo, the language of the Reef Islands? *Oceanic Linguistics* 46.456–498.
- Sasse, Hans-Jürgen. 1985. Sprachkontakt und Sprachwandel: Die Gräzisierung der albanischen Mundarten Griechenlands. *Papiere zur Linguistik* 32.37–95.
- Siegel, Jeff. 2008. *The emergence of pidgin and creole languages*. Oxford: Oxford University Press.
- Thurston, William. 1989. How exoteric languages build a lexicon: esoterogeny in western New Britain. In Ray Harlow and Robin Hooper, eds *VICAL 1: Oceanic languages. Papers from the Fifth International Conference on Austronesian Linguistics, part 2*, 555–579. Auckland: Linguistic Society of New Zealand.
- — 1994. Renovation and innovation in the languages of north-western New Britain. In T. Dutton and D. Tryon, eds *Language contact and change in the Austronesian world*, 573–609. Berlin: Mouton de Gruyter.
- Tryon, Darrell. 1994. Language contact and contact-induced language change in the eastern outer islands, Solomon Islands. In T. Dutton and D. Tryon, eds *Language contact and change in the Austronesian world*, 611–648. Berlin: Mouton de Gruyter.
- — 2002. Buma. In J. Lynch, M. Ross and T. Crowley, eds *The Oceanic languages*, 573–586. London: Curzon.
- Tryon, Darrell and Brian D. Hackman. 1983. *Solomon Islands languages: an internal classification*. Canberra: Pacific Linguistics.
- Tryon, Darrell, ed. 1995. *Comparative Austronesian dictionary. An introduction to Austronesian studies*. Berlin/New York: Mouton de Gruyter.
- Tua, Benjamin and Peter Lincoln. 1979. *Our custom way Vanikoro: the custom book*. Auckland: University of Auckland.
- Wurm, Stephen A. 1976. The Reef Islands–Santa Cruz family. In S.A. Wurm, ed. *New Guinea area languages and language study, vol. 2: Austronesian languages*, 637–674. Canberra: Pacific Linguistics.